THE UNIVERSITY *of* EDINBURGH

THE ROLE OF SUMO MODIFICATION IN HEPATOCYTE DIFFERENTIATION

Zara Hannoun
BSc Honours

University of Edinburgh
2011

This dissertation is submitted for the degree of Doctor of Philosophy

TABLE OF CONTENTS

CHAPTER TWO: MATERIALS AND METHODS

CHAPTER FOUR: THE IMPORTANCE OF HEPATOCYTE NUCLEAR FACTOR 4 α HEPATIC ENDODERM DIFFERENTIATION

CHAPTER FIVE: DEFINING THE RELATIONSHIP BETWEEN HEPATOCYTE NUCLEAR FACTOR 4 α AND SUMOYLATION IN HEPATIC ENDODERM DIFFERENTIATION

ABSTRACT

Primary human hepatocytes are a scarce resource with variable function, which diminishes with time in culture. As a consequence their use in tissue modelling and therapy is restricted. Human embryonic stem cells (hESCs) could provide a stable source of human tissue due to their properties of self-renewal and their ability to give rise to all three germ layers. hESCs have the potential to provide an unlimited supply of hepatic endoderm (HE) which could offer efficient tools for drug discovery, disease modelling and therapeutic applications. In order to create a suitable environment to enhance HE formation, hESC culture needed to be standardised. As such, a media trail was carried out to define serum free media capable of maintaining hESC in a pluripotent undifferentiated state. We also ensured hESC cultured in the various media could be directly differentiated to HE in a reproducible and efficient manner. The project then focused on the effect of post-translational modifications (PTMs), specifically SUMOylation, in hepatocyte differentiation and its subsequent manipulation to enhance HE viability. SUMOylation is a PTM known to modify a large number of proteins that play a role in various cellular processes including: cell cycle regulation, gene transcription, differentiation and cellular localisation.

We hypothesised that SUMO modification may not only regulate hESC self renewal, but also maybe required for efficient hESC differentiation. We therefore interrogated the role of SUMOylation in hESC differentiation to hepatic endoderm (HE). hESC were differentiated and the cellular lysates were analysed by Western blotting for key proteins which modulate the conjugation and de conjugation of SUMO. We demonstrate that peak levels of SUMOylation were detectable in hESC populations

and during cellular differentiation to definitive endoderm (DE), day 5. Following commitment to DE we observed a decrease in the level of SUMO modified proteins during cellular specialisation to a hepatic fate, corresponding with an increase in SENP 1, a SUMO deconjugation enzyme. We also detected reduced levels of hepatocyte nuclear factor 4 α (HNF4α), a critical regulator of hepatic status and metabolic function, as SUMOylation decreased. As a result, we investigated if HNF4α was SUMOylated and if this process was involved in modulating HNF4α's critical role in HE.

HNF4α is an important transcription factor involved in liver organogenesis during development and is a key regulator for efficient adult liver metabolic functions. We observed a decreasing pattern of HNF4α expression at day 17 of our differentiation protocol in conjunction with a decrease in SUMO modified proteins. In order to further investigate and validate a role of SUMOylation on HNF4α stability Immunoprecipitation (IP) was employed. HNF4α protein was pulled down and probed for SUMO 2. Results show an increase in the levels of SUMO2 modification as the levels of HNF4α decrease. Through deletion and mutation analysis we demonstrated that SUMO modification of HNF4α was restricted to the C-terminus on lysine 365. Protein degradation via the proteasome was responsible for the decrease in HNF4α, demonstrated by the use of a proteasome 26S inhibitor MG132. Additionally, a group at the University of Dundee has shown that polySUMOylation of promyelocytic leukaemia protein (PML) leads to its subsequent ubiquitination via RNF4, an ubiquitin E3 ligase, driving its degradation. Using an in vitro ubiquitination assay, we show that polySUMOylated HNF4α is preferentially

ubiquitinated in the presence of RNF4. Overall polySUMOylation of HNF4α may reduce its stability by driving its degradation, hence regulating protein activity.

In conclusion, polySUMOylation of HNF4α is associated with its stability. HNF4α is subsequently important for HE differentiation both driving the formation of the hepatocytes and in maintaining a mature phenotype, in agreement with a number of different laboratories. Creating the ideal environment for sustaining mature functional hepatocytes, primary and those derived from hESCs and iPSCs, is essential for further use in applications such as drug screening, disease modelling and extracorporeal devices.

DECLARATION

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where indicated in the text.

This thesis includes material that has been the outcome of work done in collaboration. The portions of the thesis that were done in collaboration have been explicitly designated within the Acknowledgements. The names of the collaborators and the nature of their contribution have been stated.

The work in this thesis has not been submitted for any other degree or professional qualification.

FIGURES AND TABLES

LIST OF FIGURES AND TABLES

ABBREVIATIONS

AA	Activin A
AD	Activation Domain
ADP	Adenosine Diphosphate
AFP	Alpha Feto Protein
ALB	Albumin
AMP	Adenosine Monophosphate
APOF	Apolipoprotein F
ASGPR	Asialogly co-protein Receptor
ATP	Adenosine Triphosphate
BAL	Bio-artificial Liver
BCA	Bicinchoninic Acid
BD	Binding Domain
BEC	Biliary Epithelial Cells
BMP	Bone Morphogenic Protein
BSA	Bovine Serum Albumin
C/EBP	CCAAT - Enhancer Binding Protein
$CaCl_2$	Calcium Chloride
CBP	cAMP Response Element Binding Protein
CM	Conditioned Media
CMV	Cytomegalovirus
CO_2	Carbon Dioxide
CT	Carboxyl Terminus
Cyp	Cytochrome P450
DAB	Diaminobenzidine
DMEM	Dulbecco's Modified Eagle's medium
DMSO	Dimethylsufoxide
DR	Direct Response Elements
DRP	GTPase Dynamin Related Protein
DS	Double Stranded
DTT	Dithiothreitol
EB	Embryoid Body
ECL	Enhanced Chemiluminescence
ECM	Extra-cellular Matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal Growth Factor
EHS	Engelbreth-Holm-Swarm
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
FOX	Forkhead Box
GAP	GTPase Activating Protein
GMP	Good Manufacturing Practice
GOI	Gene of Interest

GRGDS	Glycine - Arginie - Aspartic Acid - Serine
GRIP	Glucocorticoid Receptor Interacting Protein
HBV	Hepatitis B Virus
hCCM	Human Cell Conditioned Media
HE	Hepatic Endoderm
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hESCs	Human Embryonic Stem Cells
HGF	Hepatocyte Growth Factor
HIF	Hypoxia Inducible Factor
His	Histidine
HNF	Hepatocyte Nuclear Factor
HRP	Horse Radish Peroxidase
HSF	Heat Shock Factor
ICA	Islet Cell Auto Antigen
Id	Inhibitor of Differentiation Genes
IHC	Immunohistochemistry
IP	Immunoprecipitation
iPSCs	Induced Pluripotent Stem Cells
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KCL	Potassium Chloride
KD	Knock Down
KOSR	Knockout Serum Replacement
L15	Leibovitz's L-15 Media
LB	Luria Bertani Broth
LBD	Ligand Binding Domain
LDL	Low Density Lipoproteins
LDS	Lithium Dodecyl Sulphate
LIF	Leukaemia Inhibitory Factor
LTR	Long Terminal Repeat
LV	Lentiviral Vector
MAPK	RAS/MAP Kinase Pathway
MEFs	Mouse Embryonic Fibroblasts
MEM	Minimum Essential Medium
MES	2-(N-morpholino)ethanesulfonic acid
$MgCl_2$	Magnesium Chloride
MODY	Maturity Onset Diabetes of the Young
MOI	Multiplicity of Infection
MT	mTeSR-1®
NaCl	Sodium Chloride
NADH	Nicotinamide Adenine Dinucleotide
NEAA	Non Essential Amino Acids
NGF	Nerve Growth Factor
NMDA	N-Methyl-D-Aspartate
NP-40	Nonidet-P40
NT	Amino Terminus
Nup	Nuclear Pore Protein

OC	One-cut
OE	Over Expression
OSM	Oncostatin M
PAS	Peptide Acrylate Surface
PBS	Phosphate Buffered Saline
PBST	Tween Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PET	Polyethylene Terephthalate
PFA	Paraformaldehyde
PHH	Primary Human Hepatocytes
PI3K	Phosphoinositide 3-Kinase
PIAS	Protein Inhibitor of Activated STAT
PKA/C	Protein Kinase A/C
PM	Point Mutant
PMEDSAH	2-(methacryloyloxy) ethyl dimethyl-(3-sulfopropyl) ammonium hydroxide
PML	Promyelocytic Leukaemia
PMSF	Phenylmethanesulfonylfluoride
PRC	Polycomb Repression Complex
PTM	Post Translational Modification
PVDF	Polyvinylidene Fluoride
RD	Repression Domain
RISC	RNA Induced Silencing Complex
RPKM	Reads per Kilo Base per Million Reads
RRE	Reverse Response Element
RSV	Rous Sarcoma Virus
RT	Reverse Transcription
SAE	SUMO Activating Enzyme
SDS	Sodium Dodecyl Sulphate
SENP	Sentrin Specific Peptidase
SIM	SUMO Interaction Motif
SMA	Smooth Muscle Actin
SOC	Super Optimal Broth
SOX	Sex Determining Region Y Box
SR	Serum Replacement
SSC	Salt Sodium Citrate
SSEA	Stage Specific Embryonic Antigens
STAT	Signal Transducer and Activator of Transcription
SUMO	Small Ubiquitin Like Modifier
TAT	Tyrosine Aminotransferase
TBPA	Thyroxin Binding Pre Albumin
TGF	Tumour Growth Factor
TO	Tryptophan Oxidase
Ub	Ubiquitin
UTR	Untranslated Region
WT	Wild Type

ACKNOWLEDGMENTS

Special thank you to Dr. David Hay and Professor Sir Ian Wilmut for the opportunity to carry out my PhD and their continual guidance and support throughout the project. Further thanks is extended to Dr. Hay and Prof. Wilmut's lab members who provided good advice and constant support, most particularly Ms. Judy Fletcher for her unwavering conviction of my abilities. A thank you to Dr. Ronald Hay and Mr. Ellis Jaffray for their positive input and generosity. Many thanks to Timothee Cezard and the staff at the GenePool for their help with the SOLEXA data. Finally, a special thank you to my family for their resolute belief in me.

PUBLICATIONS

The following papers were published during the course of my study for the degree of Doctor of Philosophy. The findings of this thesis are included, in part, in these publications; however, others are part of research collaborations:

1. Deriving Hepatic Endoderm from Pluripotent Stem Cells

Claire N. Medine, Zara Hannoun, Sebastian Greenhough, Catherine M. Payne, Judy Fletcher, David C. Hay
Springer Protocols Handbook Series, 2011

2. Lineage specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development

Alexey Ruzov, Yanina Tsenkina, Andrea Serio, Tatiana Dudnakova, Judy Fletcher, Tatiana Chebotareva, Steve Pells, Zara Hannoun, Gareth Sullivan, Siddharthan Chandran, David Hay, Mark Bradley, Ian Wilmut, and Paul De Sousa. Cell Research, Nature, 2011, Article in Press

3. Unbiased screening of polymer libraries to define novel substrates for functional hepatocytes with inducible drug metabolism.

David C. Hay, Salvatore Pernagallo, Juan Jose Diaz-Mochon, Claire N. Medine, Sebastian Greenhough, Zara Hannoun, Joerg Schrader, James R. Black, Judy

Fletcher, Donna Dalgetty, Alexandra I. Thompson, Philip N. Newsome, Stuart J. Forbes, James A. Ross, Mark Bradley, John P. Iredale
Stem Cell Research, 2011, Article in press

4. The comparison between conditioned media and serum-free media in human embryonic stem cell culture and differentiation.

Hannoun Z, Fletcher J, Greenhough S, Medine C, Samuel K, Sharma R, Pryde A, Black JR, Ross JA, Wilmut I, Iredale JP, Hay DC.
Cell Reprogram. 2010 Apr;12(2):133-40.PMID: 20677928

5. Post-translational modification by SUMO.

Hannoun Z, Greenhough S, Jaffray E, Hay RT, Hay DC.
Toxicology. 2010 Dec 30;278(3):288-93. Epub 2010 Jul 30.PMID: 20674646

6. Hepatic endoderm differentiation from human embryonic stem cells.

Hannoun Z, Filippi C, Sullivan G, Hay DC, Iredale JP.
Curr Stem Cell Res Ther. 2010 Sep;5(3):233-44. Review.PMID: 20214561

7. Generation of functional human hepatic endoderm from human induced pluripotent stem cells.

Sullivan GJ, Hay DC, Park IH, Fletcher J, Hannoun Z, Payne CM, Dalgetty D, Black JR, Ross JA, Samuel K, Wang G, Daley GQ, Lee JH, Church GM, Forbes SJ, Iredale JP, Wilmut I.
Hepatology. 2010 Jan;51(1):329-35.PMID: 19877180

8. Highly efficient differentiation of hESCs to functional hepatic endoderm requires ActivinA and Wnt3a signalling.

Hay DC, Fletcher J, Payne C, Terrace JD, Gallagher RC, Snoeys J, Black JR, Wojtacha D, Samuel K, Hannoun Z, Pryde A, Filippi C, Currie IS, Forbes SJ, Ross JA, Newsome PN, Iredale JP.
Proc Natl Acad Sci U S A. 2008 Aug 26;105(34):12301-6. Epub 2008 Aug 21.PMID: 18719101

9. The inhibitory role of stromal cell mesenchyme on human embryonic stem cell hepatocyte differentiation is overcome by Wnt3a treatment.

Fletcher J, Cui W, Samuel K, Black JR, Hannoun Z, Currie IS, Terrace JD, Payne C, Filippi C, Newsome P, Forbes SJ, Ross JA, Iredale JP, Hay DC.
Cloning Stem Cells. 2008 Sep;10(3):331-9.PMID: 18479212

CHAPTER ONE

INTRODUCTION

CHAPTER 1

INTRODUCTION

The focus of this thesis was directed towards studying human liver biology. This was done by creating a standardised model for investigating the mechanisms regulating human embryonic stem cell differentiation into hepatocytes. In addition, the roles of post translational modification, specifically SUMOylation, were investigated within hepatic differentiation, focusing on the effect on hepatocyte nuclear factor 4 α (HNF4 α) regulation, a vital transcription factor for hepatic development, differentiation and metabolism.

1.1 HUMAN EMBRYONIC STEM CELLS

1.1.1 CHARACTERISTICS OF HUMAN EMBRYONIC STEM CELLS

Human embryonic stem cells (hESCs) are isolated from the inner cells mass of blastocyst stage embryos (Thomson et al., 1998). hESCs possess two unique properties; self renewal, capable of dividing indefinitely in culture, and pluripotency, the ability to differentiate to cell types from all three germ layers; the endoderm, mesoderm and ectoderm, Figure 1.1 (Cai et al., 2006, Hoffman et al., 2005, Vazin et al., 2010). hESCs are defined by the expression of pluripotent transcription factors; Octamer 3/4, Nanog and SOX2 in addition to their cell surface antigen pattern; defined by the presence of stage specific embryonic antigens (SSEA) 3, 4, Tra 1-60 and 1-81(Pera et al., 2000, Hoffman et al., 2005). Pluripotency is measured using a combination of spontaneous (Fletcher et al., 2008) and directed differentiation to cell

types, including cardiomyocytes (Melkoumian et al., 2010), hepatocytes (Hay et al., 2008), neuronal derivatives (Lee et al., 2007) and muscle (Xiao et al., 2008).

1.1.2 HUMAN EMBRYONIC STEM CELL CULTURE DEFINITION

In order to realise the full potential of hESCs, it is important to maintain cells in defined and optimised conditions. Creating a standardized culture system for hESC maintenance will provide an ideal resource for the study of human biology, in vitro modelling and cell based therapies. To date, the most common method of maintaining hESCs is on mouse embryonic fibroblasts (MEFs), which are capable of secreting the required growth factors to sustain replication and provide extra-cellular matrix attachment support imperative for undifferentiated regeneration (Ilic, 2006, Vazin et al., 2010). However, limitations of MEFs include xeno-contamination, lack of definition and variability between batches of cells (Barbara S Mallon et al., 2006). As a result, human feeders have been developed as a successful replacement for animal feeders in an effort to create a xeno-free hESC culture environment (Amit et al., 2003). Once again, the use of human feeders suffer from variability, limited sourcing and an impure hESC population altering the accuracy of results following an investigation (Barbara S Mallon et al., 2006). Therefore, other culture conditions have been explored.

Feeder free conditions are important if hESCs are to be employed for use in clinical or therapeutic applications. As a result, research was focused towards generating a fully defined system whereby ideally, the media are supplemented with recombinant

growth factors and the extra cellular matrix support is synthetically produced. A number of groups have

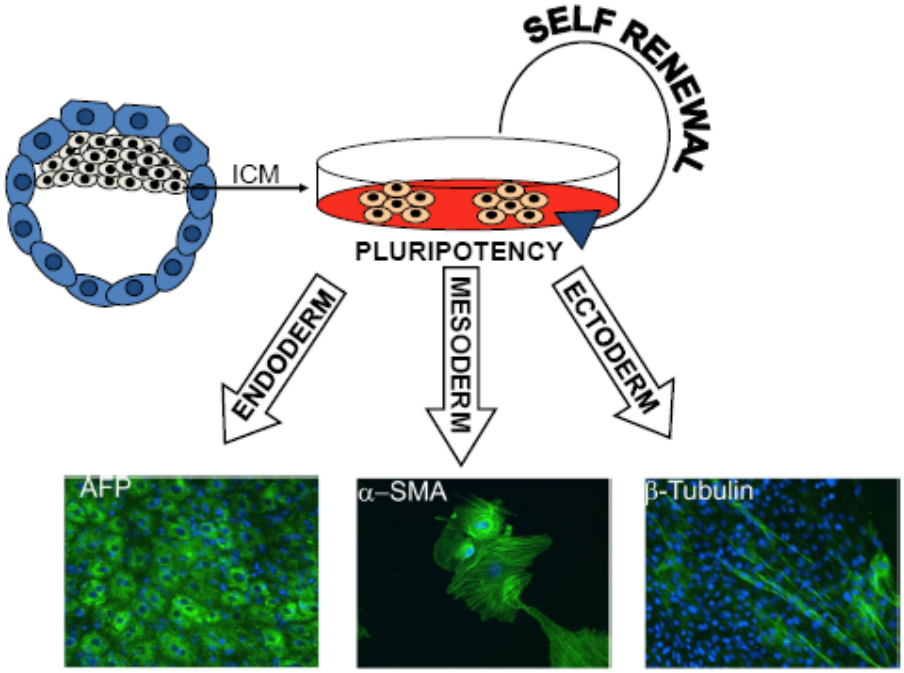


Figure 1.1 – Properties of Human Embryonic Stem Cells.
hESCs are isolated from the inner cell mass of the blastocyst stage embryo and posses two unique properties that make them ideal tools in research, clinical and therapeutic applications. hESCs are capable of indefinite self renewal, defined by unlimited cell division in culture and pluripotency, allowing the cells to form somatic cells from all three germ layers; endoderm defined by alpha feto protein (AFP) positive cells, mesoderm defined by α-smooth muscle actin (α-SMA) and ectoderm depicted by β-tubulin expressing cells.

significantly contributed to developing the ideal culture system. Xu and colleagues were the first to establish feeder free hESC culture using MEF conditioned media (CM) to maintain healthy self-renewal (Xu et al., 2001) on Matrigel™, a matrix secreted by Engelbreth-Holm-Swarm (EHS) mouse sarcoma cells (Ludwig et al., 2006). Subsequently, hESCs have been successfully maintained on recombinant vitronectin (Braam et al., 2008) and laminin (Rodin et al., 2010) resulting in improved culture definition. In conjunction, CM has been adequately replaced by a variety of defined and serum free media (Chin et al., 2007), (Chin et al., 2010, Hannoun et al., 2010). Additional research has been employed to create a

standardized completely defined hESC culture system capable of self-renewal whilst maintaining pluripotency. Clinical grade hESCs have been developed by culturing the cells in a fully xeno-defined medium, RegES (Rajala et al., 2010). The hESCs exhibited the correct morphology and gene expression associated with embryonic stem cells in addition to their ability to spontaneously differentiate into all three germ layers and be directly differentiated into neural cells and cardiomyocytes (Rajala et al., 2010). The xeno-free defined media formulation can in turn be optimised for the culture, expansion and differentiation of other cell types.

1.1.3 The Effect of the Microenvironment on hESCs

A number of differences have been noted between various cells lines relating to their default differentiation lineage and culture properties (Tavakoli et al., 2009). Therefore, it can be concluded that the derivation technique and the culture environment play a significant role in priming hESC properties, in conjunction with their varying genotypes (Frost et al., 2011, Reijo Pera et al., 2009). Studies have shown that the environmental conditions affect cellular epigenetics, which in turn alter the behaviour of the respective hESC line (Frost et al., 2011). Despite the origin of the stem cells, variations between the methylation patterns on genes were noted affecting the integrity of the imprinted gene expression (Frost et al., 2011). In addition, significant differences in epigenetic and genetic stability, transcriptional profiling and pluripotency marker expression have been documented when comparing various hESC lines (Allegrucci et al., 2007).

hESCs have a comparable gene expression pattern with epiblast cells, unlike mouse embryonic stem cells which are more comparable to the inner cell mass (Tesar et al., 2007, Brink et al., 2008). The differences in the state and expression patterns of hESCs and cells directly isolated and analysed from the inner cell mass may be attributed to culture conditions (Brink et al., 2008). In an effort to 'push back' the state of hESC to a state comparable to the inner cell mass, Hanna and colleagues induced the expression of Oct4, Klf4 and 2 in hESCs, and supplemented the media with leukaemia inhibitory factor (LIF) and glycogen synthase kinase 3 and mitogen activated protein kinase pathway inhibitors. This resulted in generating a more immature state of hESCs, which was now comparable to mouse ESCs (Hanna et al., 2010).

Studies have also demonstrated the ability to maintain the undifferentiated and pluripotent state of mouse embryonic stem cells with the use of two factors; leukaemia inhibitory factor (LIF) and bone morphogenic protein (BMP) (Ying et al., 2003). LIF maintains stem cell self-renewal by activating the STAT3 transcription factor (Ying et al., 2003). However, the cells begin to differentiate towards the neural lineage in the presence of serum, thus BMP was utilised to block differentiation by promoting the expression of the inhibitor of differentiation (Id) genes (Ying et al., 2003). Further investigation has defined fibroblast growth factor 4 (FGF4) as an essential component for driving differentiation towards the neural and mesodermal lineage in the absence of LIF (Kunath et al., 2007). Interestingly, FGF4 does not block the BMP pathway and therefore promotes differentiation, but is required to redirect the subsequent effects of BMP signalling (Kunath et al., 2007). Therefore

supporting previous studies depicting the secondary action of BMP over FGF (Stern, 2005). Culturing embryonic stem cells in optimum environmental conditions allows for the intricate control of the self-renewal and pluripotent abilities of the cells, which has yet to be achieved in hESCs.

1.2 HEPATIC ENDODERM

1.2.1 HUMAN LIVER DEVELOPMENT

Liver development is an intricate process regulated by growth factors, transcription factors, cytokines and cellular interactions, Figure 1.2. Hepatic formation during embryogenesis has been classified into three stages: induction, specification and maturation. A distinct set of growth factors coordinated with transcriptional activators are responsible for the regulation of each stage. Hepatic induction is defined by the formation of the definitive endoderm originating from the primitive streak at the 8-12 somite stage in humans, Figure 1.2. Subsequently, the foregut is generated by the invagination of the endoderm. The ventral foregut is accurately positioned from the developing cardiac mesoderm, permitting signal responses required for inducing its development towards hepatic fate, Figure 1.2. Hepatic specification occurs once the primary liver bud forms, growing outward from the ventral foregut at the 13-20 somite stage. Angiogenesis and vasculogenesis, mediated by oxygen sensing factors, occur simultaneously during liver bud formation, leading to the vascular anatomy necessary for hepatic function. Liver maturation is completed after birth, whereby hepatocyte proliferation decreases and mature function develops, Figure 1.2. (Duncan, 2003, Costa et al., 2003, Zaret, 2000)

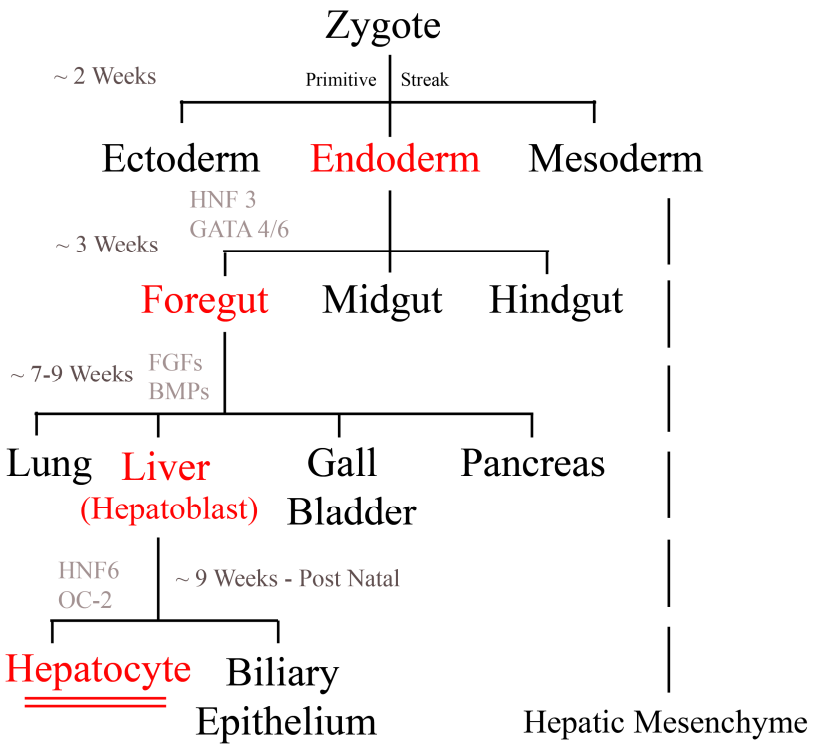


Figure 1.2 – Human Liver Development.
Liver development occurs throughout development whereby the primitive streak post embryo formation differentiates to form all three germ layers; the ectoderm, endoderm and mesoderm. The endoderm is then primed by hepatocyte nuclear factor 3 (HNF3) and GATA4/6 down the foregut lineage followed by liver specification induced by bone morphogenic proteins (BMPs) and fibroblast growth factors (FGFs). Hepatocyte nuclear factor 6 (HNF6) and one cut 2 (OC-2) transcription factors are responsible for hepatocyte commitment followed by hepatic maturation activated by hepatocyte growth factor (HGF) and Oncostatin M (OSM), which continues post natal.

1.2.1.1 Hepatic Induction

Fibroblast growth factors (FGF) 1 and 2, secreted by the cardiogenic mesoderm, and bone morphogenetic protein (BMP), produced by the septum transversum mesenchyme induce the initial phase of hepatic development, Figure 1.2 (Rossi et al., 2001). FGF signalling activates the RAS/MAP kinase pathway (MAPK) leading to enhanced hepatic gene expression and nascent hepatocyte stability. The FGF signalling pathway is capable of initiating down stream effects either through the

default MAPK pathway or through the phosphatidylinositol 3-kinase pathway, active in the foregut endoderm (Calmont et al., 2006). As liver development progresses, morphogenetic changes result in the distancing between the hepatic endoderm and the cardiac mesoderm (Jung et al., 1999). Consequently, the concentration of FGF decreases inhibiting the differentiation into an anterior fate such as the lungs. FGF concentration gradients induce varied responses vital for synchronizing developmental events (Serls et al., 2005). BMP 2 and 4 operate in accordance with FGF signalling to encourage hepatic competence and specification of the primitive endoderm via GATA4, a GATA zinc finger transcription factor, activation (Huang et al., 2008).

1.2.1.2 Hepatic Specification

The Wnt/β-Catenin pathway is also involved in hepatic induction and is responsible for stimulating a combination of differentiation and proliferation of pre-hepatic endodermal cells. This is carried out via the activation of the Frizzled receptor in response to the Wnt ligand (McLin et al., 2007). Hepatoblast proliferation is vital for liver development and hepatocyte repopulation during injury, which is stimulated by Wnt9a, secreted by sinusoidal and stellate cells in the liver (Matsumoto et al., 2008). Additionally, Wnt9a has been demonstrated to increase glycogen synthase production and enhance glycogen accumulation in the chick (Matsumoto et al., 2008). Wnt3a has been shown to promote liver lobe periphery proliferation and regulate liver morphology (Suksaweang et al., 2004). A zebrafish Wnt2b homologue is a key regulator in liver specification as well as a possible factor affecting hepatoblast proliferation (Ober et al., 2006). Wnt3a has also been documented to be

expressed during hepatoblast differentiation into hepatocytes. Its subsequent addition to a cell culture model resulted in a homogenous population of stem cell derived HE (Hay et al., 2008). β-Catenin exhibits its effects by interacting with FGF 10 stimulating hepatoblast proliferation (Jung et al., 1999).

The transforming growth factor α (TGF α) pathway is involved in differentiating hepatoblasts towards the biliary lineage (Shiojiri, 1997), which is assumed to be achieved by the formation of a TGF α signalling gradient, however, the exact mechanism is unknown. Hepatocyte nuclear factor 6 (HNF6) and One-cut 2 (OC2) transcription factors have been demonstrated to regulate Follistatin and β-2-Macroglobulin, TGF α antagonists. Studies in HNF6 knockout mice revealed an increase in the activation of the TGF α pathway, in addition to limited segregation observed between the biliary and hepatic cell lineages (Clotman et al., 2005). Therefore, HNF6 is required for repressing the expression of the tgf α receptor II gene (Clotman et al., 2005). Furthermore, Smad2 and 3, modulators of TGF α ligands, knockout mice display liver hyperplasia; which is rescued by addition of hepatocyte growth factor (HGF). It could be hypothesized that HGF regulates hepatic cellular organization and further contributes to hepatic specification. In terms of in vitro modelling, it is reasonable to suggest that inhibitors of the TGF α pathway may commit hESCs towards the hepatic lineage. On the other hand, modelling the system in 3D and the generation of a TGF α signalling gradient may lead to the specification of the two cell types, both the hepatic and biliary, thereby potentially establishing hepatic architecture.

1.2.1.3 Hepatic Maturation

HGF and Oncostatin M (OSM) are soluble molecules that promote hepatocyte maturation. HGF is secreted by an array of tissues present in the liver including the septum transversum mesenchyme, endothelial cells and hepatoblasts. HGF is responsible for binding to the c-Met receptor resulting in the activation of the SEK1/MKK4 and c-Jun signalling pathways (Kamiya et al., 2001). Activating the above signalling cascades results in the up-regulation of glucose-6-phosphate, tyrosine amino transferase, carbamoyl-phosphate synthase and albumin expression, all of which are associated with mature liver (Kamiya et al., 2001). OSM, is an IL-6 related cytokine, is produced by hematopoietic cells present in mid-fetal livers and exerts its effect by activating the signal transducer and activator of transcription (STAT3) and Ras Pathways *via* the OSM receptor (Wong et al., 2003). A number of hepatic gene promoters are directly regulated by STAT3 *via* binding sites in the 5' regulatory regions (Runge et al., 1998). Like HGF, OSM induces the expression of various liver markers (Kamiya et al., 2001). At birth OSM is down regulated whilst HGF is unregulated (Kagoshima et al., 1992). The liver undergoes a functional switch at birth accelerating liver maturation resulting in the preparation of the organ to respond to environmental changes; including a significant increase in the levels of glycogen and changes in oxygen tension (Kinoshita et al., 2002).

Developing a complete understanding of the signalling pathways involved in liver development will provide significant insight behind the underlying regulatory mechanisms. This in turn was translated into various successful differentiation protocols utilized for generating HE from hESCs.

1.2.2 SOURCES OF HEPATIC ENDODERM

Hepatocytes are in limited supply and existing sources suffer from various disadvantages, which limits both the ability to study human liver function and the applications utilising functional HE. Primary human hepatocytes isolated from human liver are the gold standard for use in research investigations and clinical applications. However, they are a scare resource due to a lack of donor livers and suffer from diminished function and viability in long-term culture (Kim et al., 2007), insufficient for high throughput drug screening and bio-artificial liver devices (BAL). Liver grafts from matching donors have proved successful in diseased liver cases, but due to limited supply, it cannot be regarded as a stable form of treatment. Porcine and other animal isolated hepatocytes have been utilized in preliminary disease modelling studies and drug testing however, they are an inaccurate substitute for human hepatocytes, as they do not accurately mimic mechanisms found in human hepatocytes (Behnia et al., 2000). Xeno-contamination and hyper-acute rejection prohibit the employment of animal hepatocytes in therapeutic applications (Butler, 1998). In addition, risk of zoonoses transmission and incompatible liver function render xeno-derived hepatocytes as an unsuitable component for BAL devices (Le Tissier et al., 1997). Hepatoblastoma cell lines such as C3As, an immortalized subclone of the hepatoma derived HepG2 cell line (Mavri-Damelin et al., 2008), have been utilized in BAL devices but exhibit reduced function when compared to primary hepatocytes and cannot be used in therapy due to their cancerous nature(Tilles et al., 2002). Additionally, there are also concerns about tumorigenic transmission into patients. In conclusion, existing sources of hepatocytes limit the

potential of the use of hepatic endoderm (HE) as unique and effective tools for research and applications in the clinic.

Human embryonic stem cell derived HE may offer a possible solution. hESCs ability to infinitely self-renew and differentiate into any cell type may in turn provide an inexhaustible supply of mature functional HE (Hay et al., 2010). hESCs and their derivatives are ideal tools for disease modelling and genetic modification of their cell surface antigen expression may result in reduced immuno-rejection when utilized in therapy (Agarwal et al., 2008, Basma et al., 2009, Shirahashi et al., 2004). Standardised culture systems for hESC maintenance permit scalable production of HE required for high throughput drug screens and employment in BAL devices (Sharma et al., 2010). hESC directed differentiation into HE has been thoroughly examined and documented. Most techniques mimic the developmental mechanisms activated during embryogenesis.

1.2.3 Application of Human Hepatic Endoderm

Mature functional hepatocytes are in high demand for both industrial and clinical purposes. Hepatic endoderm may provide a possible long-term solution for liver disease due to the limited availability of donor livers and the lack of successful drugs on the market. To date, hepatocytes have been utilized for drug discovery investigations reducing the requirement for animal studies and increasing the safety of human clinical trials (Hay et al., 2010). Extra corporeal devices are currently in the design phase and suitable supplies of hepatocytes are being sourced (Fiegel et al., 2008). Extra corporeal devices mimic the functional liver, which in turn alleviates

the pressure on the diseased liver providing sufficient time for self-recovery (Fiegel et al., 2008). Hepatocytes could be used for cellular therapy (Alison et al., 2007, Asahina et al., 2006), on the condition they are in abundant supply and maintain the functions and properties displayed in their native environment.

1.2.4 Directed Differentiation of hESCs into Hepatic Endoderm

A vast amount of research was focused on developing effective methods for deriving functional hepatic endoderm from hESCs. The resulting HE is characterised using specific criteria including hepatic specific gene expression, serum protein production, cytochrome p450 activity, urease function and glycogen storage (Snykers et al., 2009). However, the exact combination of techniques employed varies between laboratory groups.

There are a large number of protocols for HE formation ranging from embryoid body (EB) derivation in fetal bovine serum (FBS) (Rambhatla et al., 2003), to differentiating hESCs utilizing 2-D systems using collagen, gelatin or Matrigel™ as the basement membrane. Agarwal and colleagues cultured hESCs on collagen in the presence of FBS, knockout serum replacement (KOSR) and bovine serum albumin (BSA) supplemented with Activin A (AA), fibroblast growth factor (FGF-2), hepatocyte growth factor (HGF), Oncostatin M (OSM) and dexamethasone. The resulting HE expressed a number of hepatocyte specific markers including albumin, alpha-fetoprotein (AFP), CYP3A4, CYP7A1 and was capable of glycogen storage and albumin secretion (Agarwal et al., 2008). Schwartz and colleagues generated HE solely in the presence of FGF-2 and HGF on collagen, the resulting hepatocyte like

cells expressed GATA4 and HNF1 α and 3β; were capable of albumin and urea production and exhibited Cytochrome P450 activity after treatment with Phenobarbital (Schwarts, 2005). HE has also been formed on collagen in the presence of FBS supplemented with insulin, dexamethasone, transferrin and selenious acid, where expression of albumin, transthyretin and albumin serum protein production was observed (Shirahashi et al., 2004). hESCs cultured on gelatin in the presence of HGF and nerve growth factor (NGF) generated HE expressing various hepatocyte specific genes, but showed no further function (Kuai et al., 2003). Hay and colleagues demonstrated that treatment with either Activin A/Wnt3a or sodium butyrate followed by treatment with dimethylsulfoxide (DMSO) generated immature hepatocytes. The hepatocytes could then be matured with HGF, insulin, OSM and hydrocortisone to give HE that expressed the majority of hepatocyte specific genes, were capable of glycogen storage, produced significant levels of hepatic serum proteins and had inducible cytochrome P450 activity (Hay et al., 2008, Fletcher et al., 2008, Hay et al., 2007, Hay et al., 2008).

Basma and colleagues established an efficient protocol for the purification of a hepatocyte population from a heterogeneous endodermal population. The hESCs derived EBs were plated onto Matrigel™ and treated with Activin A and FGF 2. The cells were then placed into defined media supplemented with HGF followed by dexamethasone. The resulting HE was further enriched by FACS sorting for ASGPR positive cells, a specific feature of mature hepatocytes. The pure population of HE expressed hepatic gene function comparable to adult hepatocytes (Basma et al., 2009). Cai and colleagues developed a physiological protocol that mimicked the *in*

vivo situation. This involves priming hESCs with Activin A to direct them towards definitive endoderm, followed by BMP and FGF generating hepatic endoderm. The resulting HE was matured using HGF, OSM and dexamethasone. The HE expressed a range of mature hepatic genes; however there was no expression of AFP. The HE produced significant albumin and interestingly was susceptible to infection by the hepatitis virus (Cai et al., 2007).

It has been debated that transferring culturing conditions into 3-D environments may enhance HE function; as this potentially mimics *in vivo* development more accurately. It has been proposed that culturing hepatocytes between double layers of ECM in 3-D structures will establish polarity and enhance hepatic function and viability mimicking the *in vivo* situation. One group has successfully differentiated hESCs into HE in a 3-D environment. Baharvand et al., cultured hESCs in self-renewing conditions and by using the hanging drop method, formed EBs. These EBs were then seeded onto collagen coated 3-D scaffold in culture medium supplemented with FGF-2, HGF, OSM, insulin, dexamethasone, transferrin and selenium. The resulting HE, differentiated from the EB's, expressed a number of hepatic specific genes and produced significant levels of both urea and albumin (Baharvand et al., 2006). Du and colleagues successfully constructed an ECM free synthetic culture by sandwiching a hepatocyte monolayer between two membrane like structures. The top support system consisted of a glycine - arginine- aspartic acid - serine (GRGDS) modified polyethylene terephthalate (PET) membrane. The bottom substratum consisted of a galactosylated PET membrane. This resulted in hepatic polarity including biliary excretion and enhanced function when compared to 2-D collagen

coated cultures using HE derived from hESCs (Du et al., 2008). Ng and colleagues successfully combined methylated and galactosylated collagen nanofibres that optimized the interactions required for the maintenance of functional hepatocytes. This enhanced interactions between the nanofibres and the asialoglycoprotein receptor (ASGPR), hence promoting hepatic function (Ng et al., 2005, Lu et al., 2003, Yin et al., 2003). An unbiased synthetic approach was additionally employed to isolate polyurethane and polyacrylates polymers capable of culturing and differentiating hESCs derived HE in 3D (Hay et al., 2010). Polymer 134 was demonstrated to successfully maintain differentiating hESC, where improved hepatic endoderm viability, hepatic gene expression and hepatic function, specifically enhancing drug inducible metabolism, were noted (Hay et al., 2010). The polymer was then coated onto the matrix of the bioartificial liver and supported long-term maintenance and function of the hepatocytes (Hay et al., 2010). These results signify the importance of ECM interactions for maintaining hepatic function. 2-D culture systems have been successfully employed to differentiate hESCs into functional HE; however, high fidelity hepatocyte function and improved cellular viability will inevitably be attained in a 3D system.

1.3 Liver Enriched Transcription Factors

1.3.1 The Importance of Understanding Liver Transcriptional Mechanisms

Growth factors and extracellular matrices are components important for the regulation of liver development and differentiation (Duncan, 2003, Zaret, 2000). Another important part of hepatic regulation is the array of liver enriched

transcription factors (Cereghini, 1996). The function of these transcription factors range from regulating hepatic commitment to inhibiting the differentiation into other lineages such as biliary or pancreatic (Lee et al., 2005, Watt et al., 2007). Therefore, it is essential to understand the liver regulatory transcriptional machinery in order to provide high fidelity and stable human liver models. The next section reviews important transcription factors required for liver specification, commitment and maturation, with a focus on HNF4 α.

1.3.2 LIVER ENRICHED TRANSCRIPTION FACTORS

Forkhead Box (FOX) A, GATA4 and CCAAT-enhancer binding protein (C/EBP) α transcription factors are vital for hepatic specification during liver development, specifically through the activation of the albumin gene (Bossard et al., 1998). FOXA and GATA4 relax the chromatin around the promoter region of the albumin gene allowing access and binding of other factors onto the DNA resulting in the subsequent activation of albumin expression (Cirillo et al., 2002). This type of facilitative interaction is termed 'competence' as it primes a cell of an unspecified fate down a particular lineage by initiating its ability to respond to a specific set of signals.

The forkhead protein, FOXA1, also known as hepatocyte nuclear factor HNF3 α, is a transcriptional activator of a large number of hepatocyte specific genes. FOXA1 binds to *cis* regulatory elements within the promoter regions of the albumin, α feto protein (AFP), transthyretin, tyrosine aminotransferase and PEPCK genes, triggering their expression (Navas et al., 2000). HNF1α has also been demonstrated to be

important for stimulating liver formation (Lokmane et al., 2008). Interestingly, HNF1 α null mouse models display a lack of endoderm competence and are thus unable to initiate hepatic specification.

GATA4 is a member of the GATA zinc finger transcription factors, which is involved in the regulation of the heart tube and foregut formation during embryogenesis, more specifically linked to myocardial differentiation (Charron et al., 2001). The GATA transcription factors recognize a consensus sequence AGATAG located in the promoter regions of a number of genes such as the insulin growth factor I gene (Dame et al., 2004). It has been demonstrated that GATA4 is essential for hepatic specification as deficient mouse ES cells failed to differentiate into definitive endoderm (Watt et al., 2007). As a result, GATA4 is essential for the differentiation of extra embryonic endoderm in developing embryos. In addition, GATA4 plays a vital role in pancreatic and hepatic formation in conjunction with GATA 6, FOXA1 and C/EBP.

C/EBP α is a member of the bZIP transcription factors that recognize the consensus CCAAT as either a homo- or heterodimers with other members of the family (Hattori et al., 2007). C/EBP α is a regulator of gluconeogenesis capable of activating phosphoenolpyruvate carboxykinase, a key enzyme responsible for converting oxaloacetate into phosphoenolpyruvate and carbon dioxide. Unlike FOXA1, HNF1 α and GATA transcription factors, C/EBP proteins modulate hepatocyte metabolism.

The hepatocyte nuclear factor family of transcription factors are thought to constitute the majority of liver enriched transcription factors responsible for normal liver development. HNF1 α and β, HNF3 α and β, HNF4 α and HNF6 are key components required for liver development. Each factor has a unique expression pattern, which has helped to differentiate between their individual roles (Cereghini, 1996, Costa et al., 2003, Kyrmizi et al., , Lemaigre et al., 2004). HNF1 α is exclusively expressed in fetal and hepatocytes suggesting a role in specification (Lemaigre et al., 2004, Odom et al., 2004). HNF4 α on the other hand is expressed in both fetal and adult hepatocytes (Lemaigre et al., 2004, Sladek et al., 1990). Therefore HNF4 α can be implicated in both the differentiation and maintenance of the hepatic phenotype. HNF6 is expressed in both fetal and adult hepatocytes and in biliary epithelial cells; however, its expression is ablated on biliary maturation (Lemaigre et al., 2004, Bezerra, 1998). HNF6 is central to regulation of gluconeogenic, glycolytic, bile acid synthesis pathways and is essential for hepatocyte proliferation (Jacquemin et al., 2003, Costa et al., 2003). HNF3 α and β is expressed in both fetal and adult biliary epithelial cells (BEC) and fetal hepatocytes; but expression is obliterated during maturation (Kaestner, 2000, Lemaigre et al., 2004). HNF4 α, HNF6, HNF1 α and β work synergistically and cooperatively to coordinate events throughout hepatocyte differentiation. HNF4 α and 6 co-regulate glucose-6-phosphate expression whilst HNF1 and 4 α control glucose, lipid and amino acid metabolism (Costa et al., 2003, Lokmane et al., 2008). HNF1 α is essential for the regulation of bile acid production and fatty acid oxidation within the liver (Coffinier et al., 2002). In conclusion, the HNF family are intricately involved in regulating hepatic development, differentiation and function. However, HNF4 α seems to play a key role in all the

above processes. As a result, we focused on understanding the properties and characteristics of HNF4 α.

1.3.3 HEPATOCYTE NUCLEAR FACTOR 4 α

1.3.3.1 THE PROPERTIES AND STRUCTURE OF HEPATOCYTE NUCLEAR FACTOR 4 α

Hepatocyte nuclear factor 4 α (HNF4 α) is a highly conserved member of the nuclear receptor superfamily of transcription factors (Hong et al., 2003) and is expressed in the liver, pancreas, kidney, colon and intestine (Sladek et al., 1990). HNF4 α is a constitutively active orphan receptor that binds DNA strictly as homodimer to the direct response-1 (DR1) elements with the AGGTCAGGGG(T/A)CA consensus sequence (Sladek et al., 2000). HNF4 α has also been demonstrated to bind to DR2 sequences but with a much lower affinity when compared to DR1 sequences. In addition, HNF4 α has been demonstrated to coordinate with other transcriptions factors such as C/EBP α, p300 and HNF6 to activate the expression of a number of liver specific genes (Viollet et al., 1997). Studies have shown that HNF4 α directly regulates a variety of hepatic specific genes involved in glucose, amino acid and lipid metabolism, cell structure, liver differentiation and immune function (Wisely et al., 2002).

HNF4 α is comprised of five functional regions each attributing towards a specific function carried out by the protein, Figure 1.3 (Hadzopoulou-Cladaras et al., 1997). Two activation domains are present at the amino and carboxyl terminals on HNF4 α and are responsible for promoting transcriptional activation of target genes, Figure 1.3 A (Hadzopoulou-Cladaras et al., 1997). HNF4 α directly binds to DNA via the

DNA binding domain and contains a ligand binding domain that has recently been shown to bind linoleic acid in a reversible manner (Dhe-Paganon et al., 2002, Wisely et al., 2002). The repressor domain provides another method of functional regulation responsible for inhibiting gene transcription, Figure 1.3 A. The deletion of the repressor domain actively results in the up regulation of transcriptional activation (Ihara et al., 2005). The DNA binding domain is created from two zinc fingers with four cysteine residues each linked by a hinge region, Figure 1.3 B. The ligand binding domain consists of 12 alpha helices arranged in a manner exposing a hydrophobic pocket demonstrated to bind fatty acids with high affinity in bacterial expressed HNF4 α (Dhe-Paganon et al., 2002, Wisely et al., 2002). Mammalian expressed HNF4 α only binds to linoleic acid however, no difference in transactivation function was noted (Yuan et al., 2009). HNF4 α can adopt two conformational states; an open form where the α 12 helix is extended and parallel to α helix 10 and a closed form where the α 12 helix is folded upwards (Dhe-Paganon et al., 2002). However, both the conformations allows fatty acid ligand binding, whereby the carboxylic acid group of the fatty acid binds with the guanadinium group of arginine 226 and the aliphatic chain is arranged along tunnel of hydrophobic residues (Dhe-Paganon et al., 2002).

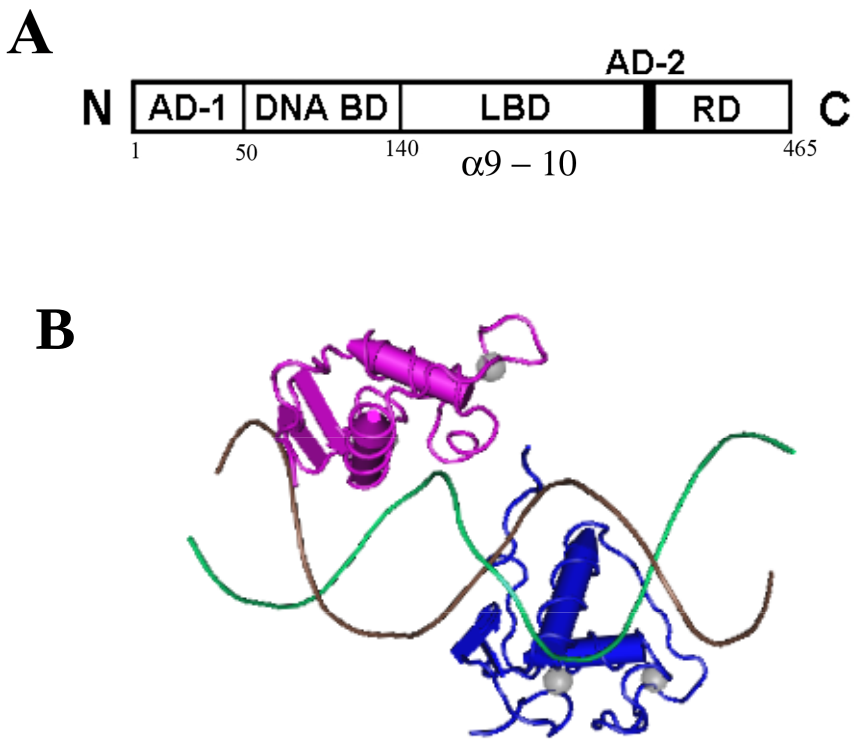


Figure 1.3 – Structure and 3D Schematic of Hepatocyte Nuclear Factor 4 α.
A. HNF4 α has 5 functional groups, including the activation domain (AD), the ligand binding domain (LBD), the DNA binding domain (DNA BD) and the repression domain (RD). (Bogan et al., 2000, Sladek et al., 1990) **B.** The 3D schematic of the HNF4 α homo-dimer zinc finger heterodimer binding to double stranded DNA. (Dhe-Paganon et al., 2002)

1.3.3.2 HEPATOCYTE NUCLEAR FACTOR 4 α Isoforms

HNF4 α has six known splice variants, whose relative expression is regulated in a tissue specific manner. The HNF4 α gene contains two promoters P1 and P2, responsible for expressing six distinct splice variants, which differ in their activation (AD1 and AD2) and repressor domains, Figure 1.4 (Briancon et al., 2006). Differential expression of the P1 and P2 driven variants have been documented in the mouse fetal and adult liver, whereby P1 driven transcripts are most likely found in the adult liver whilst P2 variants are predominant in the fetal liver, Figure 1.4

(Tanaka et al., 2006, Sladek et al., 1990)s. Variants 1 to 3 are expressed from the P1 promoter, whilst variants 4 to 6 are driven by the P2 promoter, Figure 1.4 (Sladek et al., 2000). HNF4 α v2 is predominately expressed in the adult liver, whilst HNF4 α v4 and v5 are prevalently expressed in the adult pancreas within beta cells (Ihara et al., 2005). It has been speculated that the individual HNF4 α isoforms have specific functions, suggested by their varied spatial and temporal expression. HNF4 α v5 lacks a complete activation domain-1, which could account for the reduced transcriptional activation potential when compared to variant 1 (Ihara et al., 2005). Studies have also shown that the activation domains present on the carboxyl and amino terminals of HNF4 α interact to enhance transcription (Hadzopoulou-Cladaras et al., 1997). This could also be attributed to the limited transcriptional function of HNF4 α v5. In addition, the deletion of the activation region 2 on HNF4 α variant 1 resulted in the absolute inhibition of transcriptional activity (Hadzopoulou-Cladaras et al., 1997).

Additionally, transcriptional regulation of the HNF4 α variants has been shown to be influenced by specific activators. HNF4 α v1 transcriptional activation is significantly enhanced in the presence of co activators glucocorticoid receptor interacting protein 1 (GRIP-1) and cAMP response element binding protein (CBP) (Torres-Padilla et al., 2002). On the other hand, HNF4 α v4 lacks the complete activation domain and has limited transcriptional activity, as observed with variant 5 (Torres-Padilla et al., 2002). However, HNF4 α v4 is also capable of interacting with GRIP-1 and p300 via the second activation domain but cannot do so synergistically required for enhanced transcriptional activity as observed with HNF4 α v1 (Torres-

Padilla et al., 2002). Therefore, there seem to be distinct roles for the individual variants demonstrated by their varied functional abilities and interaction capabilities with co-factors.

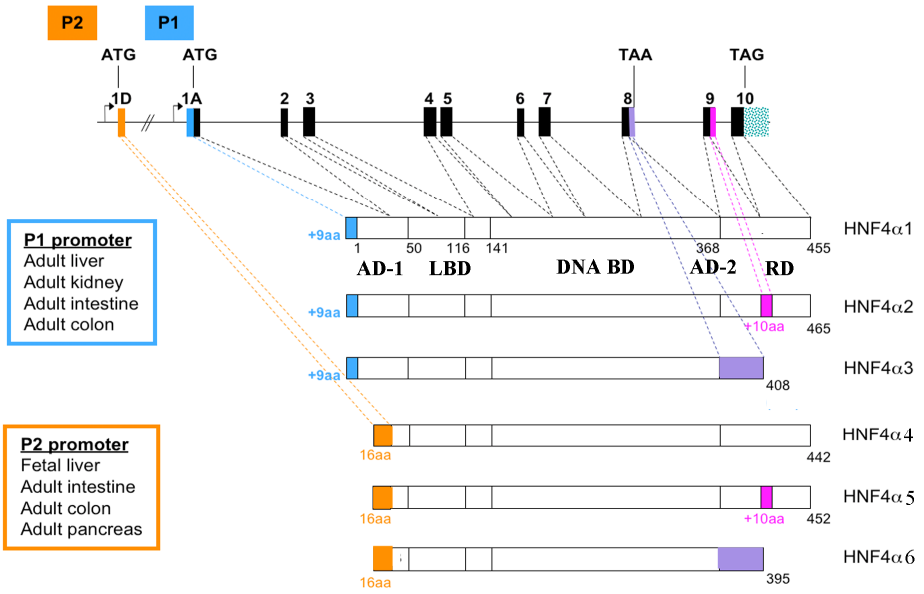


Figure 1.4 - Splicing Differences between the Various HNF4 α Variants.
HNF4 α has six different splice variants generated from two promoters, P1 and P2. All variants contain the ligand and DNA binding domains and differ in their activation and repression domains. Variants 1, 2 and 3 rely on promoter one for expression, used in the adult liver whilst variants 4, 5 and 6 rely on promoter two, used in the fetal liver. (Sladek et al., 1990)

Differentiation expression of the P1 and P2 driven HNF4 α variants has been detected in tumorigenic tissues. A significant up regulation of P1 driven variants is present in intestinal metaplasia and gastric cancer, specifically in the intestinal type (Tanaka et al., 2006). The intestinal form of the cancer is defined by the transdifferentiation of the gastric epithelial cells into an intestinal cell type (Tanaka et al., 2006). Conversely, enhancing the expression of P1 driven isoforms resulted in reduced hepatocellular and renal cell carcinoma proliferation (Lazarevich et al., 2004, Lucas et al., 2005, Chiba et al., 2005). Throughout mouse development, a switch between P2 and P1 driven HNF4 α variants is required for liver maturation

and has been documented to be involved in various cancers. However, the expression of the HNF4 α splice variants within the human liver is yet to be elucidated.

1.3.3.3 THE IMPORTANCE OF HEPATOCYTE NUCLEAR FACTOR 4 α

HNF4 α is a vital factor associated with both liver development and adult liver metabolism. HNF4 α gene knock out during embryogenesis results in embryo lethality due to visceral endoderm defects (Chen et al., 1994) This can be circumvented using conditional knock outs; however this still results in the formation of embryonic livers containing large red lesions, with discontinuous parenchyma (Parviz et al., 2003). In addition, glycogen accumulation decreased and E-cadherin expression was reduced, resulting in limited cellular interactions, in HNF4 α mutant livers (Parviz et al., 2003). Forced expression, using a recombinant retrovirus, of HNF4 α within fibroblasts induced morphological changes associated with hepatocytes, such as forming polygonal shaped cells and increased expression of E-cadherin found to be localised on the cell membranes. The results indicate the significant role of HNF4 α in regulating hepatic architecture and organisation (Parviz et al., 2003). Additionally, driving the expression of HNF4 α in a dedifferentiated hepatoma cell line (H5) resulted in re-establishing cellular junctions and polarity, inducing the expression of silenced hepatic genes such as HNF1 and activating hepatic functions such as serum protein production (Spath et al., 1997, Spath et al., 1998).

A large number of genes involved in normal liver functions such as glucose regulation, lipid synthesis, fatty acid production, xenobiotic and drug metabolism

contain HNF4 α binding sequences within their promoter regions or have been indirectly associated (Sladek et al., 2000). As such, HNF4 α is an essential regulator of adult liver functional activities. Duncan and colleagues have demonstrated that HNF4 α regulates the expression of genes required for glucose entry and glycolysis, whereby pathway activation is inhibited in livers with mutations in the glucokinase gene (Stoffel et al., 1997). HNF4 α is required for the expression of the cytochrome p450 genes and Phase II enzymes necessary for drug metabolism (Gonzalez, 2008) in addition to coordinating the majority the genes associated with the apolipoprotien and lipid pathways within the liver (Hayhurst et al., 2001). Overall, HNF4 α is a unique transcriptional activator responsible for effective liver development and the efficient maintenance of liver specific metabolic activities. In can be assumed that HNF4 α plays a key role in differentiating hESCs down the hepatic lineage.

1.3.3.4 REGULATION OF HEPATOCYTE NUCLEAR FACTOR 4 α

The mechanisms behind regulating the transcription and function of HNF4 α remain obscure. However, covalent modifications of HNF4 α have been documented. HNF4 α is phosphorylated, acetylated and methylated whereby each modification results in a change in function including DNA binding, gene transactivation, cellular localisation and protein dimerization (Viollet et al., 1997, Sladek et al., 2000). HNF4 α contains a conserved protein kinase A (PKA) consensus site located within the A box of the DNA binding domain. Phosphorylation of the serine residues via PKA resulted in decreased ability to bind DNA and reduced transcription, in vitro and vivo. (Viollet et al., 1997)

Sun and colleagues established that the phosphorylation of HNF4 α via protein kinase C on the conserved serine residue residing between the two zinc fingers reduced DNA binding, transactivation, protein stability and altered nuclear localisation. Mutating the serine residue resulted in the diminution of the above functions. Treating a hepatocellular carcinoma cell lines (HepG2) with phorobol 12-myristate 13-acetate, a PKC activator, increased the levels of cytoplasmic HNF4 α in addition to decreased levels of endogenous HNF4 α due to 26S proteasome mediated degradation. Phosphorylation of other nuclear receptors demonstrated similar results indicating a possible global mechanism for the regulation of nuclear receptor function. (Sun et al., 2007)

Post-translational modifications (PTM) of HNF4 α appear to play a significant role in regulating protein function. SUMOylation is a type of PTM that regulates the activities of proteins involved in a large number of cellular metabolic functions. Therefore, SUMOylation may affect the function of HNF4 α, specifically within hESC differentiating into hepatic endoderm.

1.4 POST TRANSLATIONAL MODIFICATIONS

Post translational modifications (PTMs) occur once a protein has been translated via the ribosomes and have been found to be involved in most or all cellular activities (Walsh et al., 2005). PTMs affect cellular processes ranging from protein degradation to regulating cell division to inducing differentiation, as such play a vital role in maintaining normal cellular metabolic activities (Walsh et al., 2005). Phosphorylation is the most common PTM and is usually involved in the regulation

of the activation or inhibition of down stream signalling pathways such as the phosphoinositide 3 kinase pathway, vital for cell proliferation, differentiation and intracellular trafficking (Karpova et al., 2006, Opazo et al., 2003). Methylation is another form of a PTM, whereby a methyl ion is covalently attached to the target protein or molecule (Imhof, 2006). DNA is commonly methylated; a reaction catalysed by DNA methyltransferase, and usually results in the repression of gene expression (Macgillivray et al., 1972). Proteins have also been documented to be methylated. Histones are the most investigated methylated proteins, resulting in the epigenetic regulation of gene expression (Nakayama et al., 2001). SUMOylation is another PTM demonstrated to affect various protein functions involved in cellular activities such as protein localisation and degradation. Further details can be found in section 1.4.1.

1.4.1 SUMOYLATION

Post translational modifications involve the addition of a chemical group following protein translation (Walsh et al., 2005). PTMs are essential to a variety of cellular processes and provide another level of protein regulation, which is usually reversible. There are a large number of PTMs that take place in the cell such as: phosphorylation (Burnett et al., 1954), methylation (Grewal et al., 2004), acetylation (Glozak et al., 2005) and glycosylation (Spiro, 2002); regulating various biological activities such as transcriptional regulation (Waby et al., 2008) and protein degradation (Orford et al, 1997).

1.4.1.1 SUMO- THE SMALL UBIQUITIN LIKE MODIFIER

SUMOylation, another type of PTM, has a diverse range of effects within the living cell (Johnson, 2004). SUMO proteins are highly conserved in a large number of species and have been shown to be important in many eukaryotic cell processes (Hayashi et al, 2002) including; cell cycle regulation, transcription, cellular localisation, degradation and chromatin organisation (Müller et al, 2001, Seeler and Dejean, 2003, Verger et al, 2003). Despite the name, SUMO only shares ~18% homology with ubiquitin (Müller et al, 2001) and is approximately 11kDa in size, comparable to the 8kDa ubiquitin molecule (Müller et al, 2001). SUMO has been found to bind to the lysine residue on the following consensus sequence; ψKxE (where ψ corresponds to a large hydrophobic amino acid, K is a lysine residue, x is any amino acid and E is a glutamic acid residue) on the target protein (Rodriguez et al., 2001). Three homologues exist in mammals, SUMO -1, -2 and -3. SUMO -2 and -3 share 95% homology with each other, but only share 50% identity with SUMO-1 (Johnson, 2004). SUMO -2 and -3 have the ability to form poly SUMO chains, covalently binding to themselves via the lysine residue at the N terminus consensus motif ψKxE. SUMO-1 lacks this consensus site and as a consequence is unable to form poly chains (Kroetz, 2005) and acts as a poly SUMO chain terminator (Ulrich, 2009).

1.4.1.2 THE MECHANISM OF THE SUMO PATHWAY

1.4.1.2.1 Pathway Overview

The SUMO conjugation pathway has a lot in common with the ubiquitination pathway. Both processes involve the use of three enzymes; E1- activating enzyme, E2 – conjugating enzyme and E3 – ligase (Takahashi et al, 2001). SUMO is bound

to its target protein via an isopeptide bond formed between an ϵ- amino group on the lysine residue on the target protein and the C terminal carboxyl group on the SUMO protein (Desterro et al, 1997). The pro-form of SUMO needs to be cleaved prior to protein conjugation (Desterro et al., 1999). This is carried out by isopeptidases, also known as the SENP SUMO deconjugating enzymes (Mukhopadhyay and Dasso, 2007). The SUMO activating enzyme (E1), SAE1/2, commences the reaction process by interacting with SUMO (activated by SENP enzymes), to form a high energy thioester bond. The SUMO conjugating enzyme (E2) then binds SUMO via its cysteine residue in its active site. This intermediate provides a highly reactive species, important in the final conjugation, usually facilitated by an E3 ligase (Kroetz, 2005). SUMO E3 ligases act to either activate Ubc9 or bring Ubc9 and the target protein within close proximity of each other, thus enhancing SUMOylation (Ulrich, 2009). They can be regarded as E3 enzymes as they are able to bind to the E2 and the substrate and facilitate the formation of the bond formed between SUMO and the target protein. It has also been shown that a large number of proteins (~ 40%) can be SUMOylated without the presence of the consensus sequence (ψKxE), demonstrating differences in substrate specificity (Ulrich, 2009).

1.4.1.2.2 The Enzymes Involved

1.4.1.2.2.1 E1

Unlike the ubiquitin (Ub) E1, the SUMO E1 exists as a heterodimer; with each monomer corresponding to a particular region of the Ub E1. The SAE subunit Aos1 (SAE1) shares similarity with the N terminus of the Ub E1, whilst Uba2 (SAE2), the second component of the SAE complex, is similar to the C terminus of the Ub E1

(Johnson et al, 1997). The monomers are never found individually and hence it is assumed that they are unable to function independently (Azuma et al, 2001). The SAE complex is responsible for preparing SUMO for transfer to the SUMO conjugating enzyme, Ubc9 (Walden et al, 2003).

1.4.1.2.2.2 E2

Ubc9 is the only known SUMO conjugating enzyme, unlike the ubiquitination pathway where each E2 has a specific set of target proteins (Hayashi et al, 2002). Ubc9 contains an active site with a cysteine residue which is responsible for binding the SUMO molecule directly to the ψKxE sequence found on the target protein (Sternsdorf et al, 1999)(Rodriguez et al., 2001). Investigations have confirmed the presence of a conserved N terminal region known to interact with SUMO is required for the efficient transfer of SUMO from the E1 enzyme to E2 (Tatham et al., 2003). Interestingly, mutations within the binding region did not affect Ubc9's ability to recognise and bind SUMO to the target protein (Tatham et al., 2003).

1.4.1.2.2.3 E3

In contrast to SUMO E2s, a larger number of SUMO E3 ligases have been discovered and have been categorized into three types; the protein inhibitor of activated STAT - signal transducer and activator of transcription (PIAS) family (Hochstrasser, 2001), the nuclear pore proteins Ran binding protein 2 and nucleoporin 358 (RanBP2/Nup358) (Pichler et al, 2002) and the polycomb group protein Pc2 (Kagey et al, 2003). E3 ligases are usually substrate specific with little redundancy found within the system.

The largest group of E3 ligases are the PIAS proteins with four genes in mammals; PIAS1, PIAS3, PIASx and PIASγ (Liu et al, 1998). The PIAS E3s have a conserved region consisting of a SAP domain responsible for binding AT rich DNA sequences and an SP-RING domain which binds to Ubc9 and promotes SUMOylation (Schmidt and Müller, 2002). They also contain SUMO interaction motifs (SIMs) that are able to directly bind SUMO (Rytinki et al, 2009). It has been found that the different PIAS proteins SUMOylate distinct sets of substrates, with occasional overlap (Schmidt and Müller, 2002). The second group of E3 SUMO ligases consists of the nuclear pore protein RanBP2 (Nup358) with only one known substrate, RanGAP1, a GTPase activating protein important in nuclear transport of proteins (Nishimoto, 1999, Saitoh et al, 1997). We have also shown that Ran GAP can be SUMO modified in vitro. The final family SUMO E3 ligase identified so far is the PC2 protein part of the polycomb group (Kagey et al, 2003). Pc2 has been shown to SUMOylate the transcriptional co-repressor CtBP, localising it to the nucleus (Lin et al, 2003), and to co-localise with PcG bodies (Kagey et al, 2003).

1.4.1.3 REGULATION OF THE SUMO PATHWAY

SUMO modification is a dynamic process involving both conjugation and deconjugation enzymes. The deconjugation enzymes function by cleaving the isopeptide bond between SUMO and the modified protein (Melchior et al, 2003). There are seven isoforms of these isopeptidases; including SENP1, SENP2, SENP3, SENP6 and SENP7 (Mukhopadhyay and Dasso, 2007). The SENPs contain a Ulp domain at their C terminus responsible for cleaving the isopeptide bond and distinct

N terminal domains that regulate their cellular localisation; suggesting each SENP has a distinct set of substrates (Mukhopadhyay and Dasso, 2007). In addition to their deconjugation role, the SENPs also play an essential role in maintaining the levels of free SUMO within the cell (Ulrich, 2009). Other forms of SUMO regulation include the E3 ligases and the presence of the consensus motif on target proteins. It has previously been stated that 40% of proteins modified by SUMO do not have the typical consensus sequence; as such, this could also be regarded as another form of regulation.

1.4.1.4 THE ROLE OF SUMO CONJUGATION WITHIN THE CELL

Over the last decade a number of groups have investigated how the SUMO pathway is regulated in response to different stimuli. In response to heat shock, erythroleukemia cells induce transcription of heat shock factor 1 (HSF1). After its translation, HSF1 is phosphorylated prior to its SUMOylation, which enhances its DNA binding (Hong et al, 2001). It is also widely recognised that SUMO alters protein activity by modulating other PTMs, such as phosphorylation and ubiquitination. For example, SUMOylation of IκBα, an important factor in the inflammatory response, prevents its ubiquitination, and therefore inhibits its degradation and subsequent NF-κB activation and nuclear translocation (Desterro et al, 1998) (Hay et al., 1999). SUMO can also regulate protein activity by modulating its interactions with other macromolecules or proteins. Various models have been proposed such as; the addition of SUMO by altering protein configuration, creating a new interaction motif affecting its function (Johnson, 2004). An interesting example of interaction motifs is arsenic induced RNF4 mediated degradation of

promyelocytic leukaemia (PML) bodies. In the presence of arsenic, PML is poly SUMOylated, and following the recruitment of RNF4, an E3 Ub ligase, PML is ubiquitinated and degraded (Tatham et al, 2008).

1.4.1.5 THE IMPORTANCE OF SUMOYLATION IN DEVELOPMENT AND CELL BIOLOGY

Various studies have shown that disruption of the SUMO pathway causes abnormal cellular differentiation. Moreover, disruption of the SUMO pathway during embryogenesis may lead to embryo lethality (Nacerddine et al, 2005, Nowak and Hammerschmidt, 2006), demonstrating the requirement for SUMOylation during development. Due to the lethal nature of Ubc9 knock outs during development in a variety of species such as mouse (Nacerddine et al, 2005), zebrafish (Nowak et al, 2006) and yeast (Johnson, 2004), other experimental strategies are necessary to determine the precise role of SUMOylation. In vitro, there has been a focus on the role of SUMOylation in a number of cell types; human embryonic stem cells (hESCs) and representatives of all three germ layers, Figure 1.5. These models, although not *in vivo*, provide a good developmental surrogate.

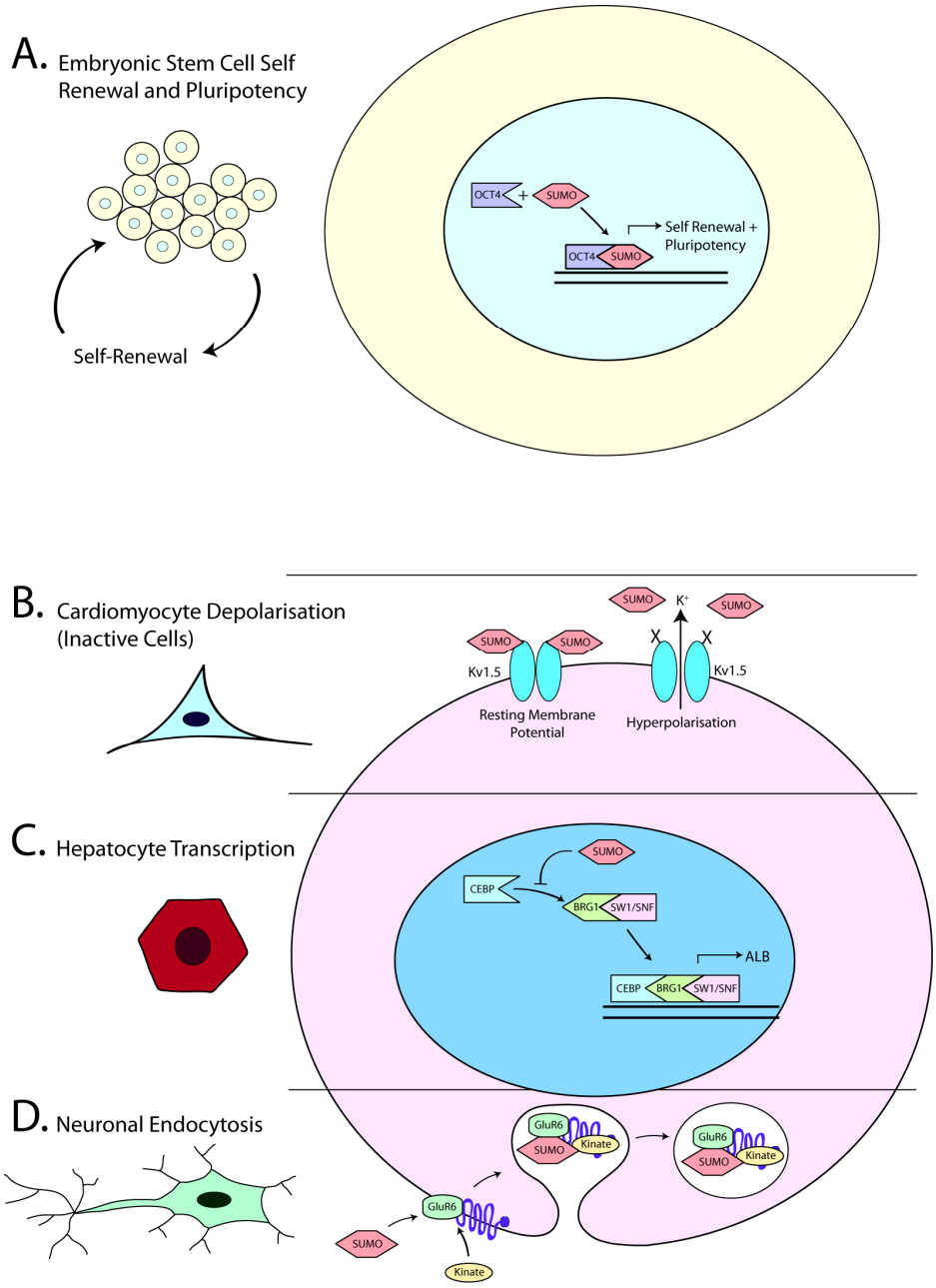


Figure 1.5- Importance of SUMOylation in Cell Biology

SUMO modification affects a number of cellular processes. (A) In human embryonic stem cells (hESCs), SUMO binds to Oct4 in the nucleus (blue), enhancing its stability and transcriptional activity, which is an important regulatory mechanism in hESC self-renewal and pluripotency. SUMO modification also plays an important role in somatic cell biology. (B) In cardiomyocytes; SUMOylation regulates the properties of the Kv1.5 potassium voltage channel located at the plasma membrane. These channels play an essential role in cardiomyocyte membrane potential. The inhibition of SUMO modification to the Kv1.5 channel results in the opening of membrane channels, exporting potassium ions, which results in cellular hyperpolarisation. (C) SUMO conjugation in hepatocytes regulates the transcriptional activity of C/EBP influencing albumin (ALB) expression within the nucleus (blue). SUMO modification of C/EBP inhibits its ability to form a complex with BRG1/SW1/SNF essential for high-level albumin expression. SW1/SNF is a chromatin-remodelling complex and BRG is a core subunit of the complex. (D) SUMOylation has also been shown to regulate the activity of the GluR6 receptor in neural cells by endocytosis. Kinate induced receptor internalisation on the cell plasma membrane is dependent on SUMO modification of GluR6, thus affecting neurone excitability.

1.4.1.5.1 Human Embryonic Stem Cells

SUMO modification has been shown to have an important role in both hESC self renewal and pluripotency (Wei et al, 2007). Oct4 is a POU transcription factor associated with the undifferentiated and pluripotent status of embryonic stem cells (Hay et al, 2004, Niwa et al, 1998). It is known to be SUMO modified, which results in its increased stability, DNA binding and transcriptional activity, Figure 1.5 A (Wei et al, 2007). Sex determining region Y box 2 (SOX2) is another important transcription factor required for embryonic stem cell self-renewal in an undifferentiated state (Zeng et al, 2006). It was recently shown by Hoof and colleagues that SOX2 is SUMO modified as a result of phosphorylation. It has been suggested that SUMO modification of SOX2 affects its transcriptional activity (Hoof et al, 2009, Hietakangas et al, 2006) but further investigation is required. The role of SUMOylation has also been determined in cell types representative of the three germ layers; endoderm, mesoderm and ectoderm.

1.4.1.5.2 Mesoderm

During development the mesoderm differentiates into muscle, cartilage, bone, blood and connective tissue (Biggers and Borland, 1976). The heart significantly relies on the coordination of various ion channels for regular function. One such voltage-gated channel is the potassium channel Kv1.5 found in atrial myocytes, which modulate membrane potential of smooth muscle cells (Lagrutta et al, 2006). Benson and colleagues have shown that Kv1.5 has two conserved consensus SUMOylation motifs, which play an important role in hyperpolarisation, Figure 1.5 A (efflux of potassium ions) (Benson et al, 2007). At the initial stages of development, the poly

comb 2 protein (Pc2), part of the polycomb repressor complex 1(PRC1), is SUMOylated. This allows efficient complex formation and its recruitment to methylated histone 3 for controlled gene silencing. On mesoderm formation, SENP2 is recruited to PRC1, deSUMOylates the Pc2 protein and allows the expression of GATA4 and 6 transcription factors essential for normal cardiac formation (Kang et al, 2010). Interestingly, in adult cardiomyocytes, SUMO modification of GATA4 results in increased transcriptional activity, and promotes cardiogenic gene activity (Wang et al, 2004).

1.4.1.5.3 Endoderm

The endoderm layer is formed during embryogenesis and is the precursor of liver, pancreas and lung amongst others (Tam et al, 2003). SUMOylation plays an important role in hepatocyte biology regulating C/EBPα, a crucial factor in hepatic differentiation (Pedersen et al, 2001, Sato et al, 2006). SUMOylation of C/EBPα prevents its association with BRG1, a core subunit in the SW1/SNF chromatin remodelling unit, leading to the inhibition of albumin expression, Figure 1.5 C (Sato et al, 2006). In line with this, it has been shown that there is a decrease in levels of SUMOylation as rat hepatocytes mature (Sato et al, 2006), suggesting an inhibitory effect of SUMOylation in hepatocyte terminal differentiation. The mitochondria are an essential component of hepatocytes, the main cell type in the liver, and are required for efficient liver function. Mitochondrial levels in the cell are dynamic and continuously undergo fusion and fission (Twig et al, 2008, Frazier et al, 2006). It has been shown that an increase in SUMO-1 expression results in an increase in mitochondrial fragmentation by stabilising the GTPase dynamin-related protein 1

(DRP1) (Harder et al, 2004). Further investigation of this pathway has revealed that SENP 5, a SUMO deconjugating enzyme, is required for normal mitochondrial morphology and levels of reactive oxidative species within the cell, partly by SUMO deconjugation of DRP1 (Zunino et al, 2007). In the pancreas, SUMO modification of islet cell auto antigen 512 (ICA512) has been shown to disrupt its binding to STAT5 and inhibit insulin and granule related gene transcription (Mziaut et al, 2006).

1.4.1.5.4 Ectoderm

Ectodermal differentiation results in the formation of the skin and nervous system (Pelton et al, 1998). SUMOylation also plays an important role in the nervous system. GluR6 is a highly expressed kinate receptor found in the brain, and is concentrated in the hippocampus (Nasu-Nishimura et al, 2010). The receptor is known to regulate neuronal excitability and as such is involved in learning, memory and synaptic plasticity (Barberis et al, 2008). It has been shown that the internalisation of the receptor upon kinate stimulation is regulated by SUMOylation. GluR6 is internalised via kinate or *N*-methyl-D-aspartate (NMDA) induced endocytosis. Only kinate induced internalisation requires GluR6 SUMOylation, Figure 1.5 C. The mutation of the SUMO consensus motif in GluR6 results in a large reduction in kinate induced GluR6 internalisation and disrupts regular synaptic function (Martin et al, 2007). Another factor important for both brain development and neuronal differentiation is MEF2A. MEF2A and its associated family members have been shown to be involved in the proliferation, differentiation and apoptosis of cells found in the developing brain (McKinsey et al, 2002). SUMOylation of

MEF2A decreases its transcriptional activity, suppressing Nur77 function; and promotes dendritic claw differentiation (Shalizi et al, 2006).

SUMOylation is an important PTM known to play roles embryonic stem cell and somatic cell biology. Given its importance in cell biology, it is critical that we understand SUMOylation in order to generate stable and high fidelity models that predict human drug toxicity. These models will not only be useful tools for toxicology, but will also provide a system whereby we can investigate the role(s) of SUMO modification in response to numerous stimuli. This will undoubtedly provide information on novel mechanisms of action with the possibility of developing new medicines and clinical intervention strategies.

1.5 THE OBJECTIVES OF THE THESIS

The thesis initially focuses on establishing a better-defined model for culturing human embryonic stem cells and their subsequent differentiation into hepatocytes. The culture system was used as a platform for studying the underlying biology of human hepatocyte differentiation from hESCs. We then focused on a key liver enriched transcription factor, hepatocyte nuclear factor 4 α (HNF4 α), and its post translational modification by SUMO. The final investigation was carried out to elucidate the relationship between HNF4 α and SUMOylation in vitro, using terminally differentiated hepatocytes.

CHAPTER TWO

MATERIALS AND METHODS

MATERIALS AND METHODS

2.1 MATERIALS AND SOLUTIONS

2.1.1 REAGENT SOLUTIONS

All reagents are supplied by Life Technologies, unless stated otherwise.

TABLE 2.1. Solutions used in a variety of techniques.

SUMO Lysis Buffer:	2% SDS 50 mM Tris pH 8 1 mM EDTA 10 mM Iodoacetimide
Protein Loading Dye:	NuPAGE® LDS Sample Buffer (x4)
SDS-PAGE Running Buffer:	NuPAGE® MES SDS Running Buffer (x20)
Western Transfer Buffer:	NuPAGE® Transfer Buffer (x20)
1 x PBS:	0.01 M Phosphate Buffer 0.0027 M Potassium chloride 0.137 M Sodium chloride
PBS Tween:	1 x PBS 0.05% (v/v) Tween-20
Coomassie Stain:	10% (v/v) Methanol (Sigma Aldrich) 20% (v/v) Acetic Acid (Sigma Aldrich) 200 mg/L Coomassie R-250
Destain Solution:	20% (v/v) Methanol (Sigma Aldrich) 10% (v/v) Glacial Acetic Acid (Sigma Aldrich) Deionised Water
LB media:	1% (w/v) Yeast Extract (Sigma Aldrich) 1% (w/v) Bacto-tryptone (Sigma Aldrich) 86 mM NaCl (Sigma Aldrich)
Native Lysis Buffer:	10% Glycerol (Sigma Aldrich) 5 M NaCl 10 mM Tris pH 7.8 5 mM DTT
Low Salt Buffer:	10 mM HEPES pH 7.9 (Sigma Aldrich) 10 mM KCL

	1.5 mM $MgCl_2$ 0.5 mM DTT
CB100:	10% Glycerol (Sigma Aldrich) 25 mM HEPES pH 7.9 0.2 mM EDTA 100 mM KCL 5 mM DTT 0.5 mM PMSF 20 mM Imidazole
CB1000:	10% Glycerol (Sigma Aldrich) 25 mM HEPES pH 7.9 0.2 mM EDTA 1000 mM KCL 5 mM DTT 0.5 mM PMSF 20 mM Imidazole
Dialysis Buffer:	20% Glycerol (Sigma Aldrich) 50 mM Tris pH 7.5 5 mM $MgCl_2$ 5 mM DTT
Elution Buffer:	20% Glycerol (Sigma Aldrich) 25 mM HEPES pH 7.9 0.2 mM EDTA 100 mM KCL 5 mM DTT 0.5 mM PMSF 200 mM Imidazole
NP40 Lysis Buffer:	50 mM Tris pH 8 150 mM NaCl 5 mM EDTA 2 mM DTT 1% NP40

2.1.2 CELL CULTURE MEDIA

TABLE 2.2. Cell culture media used for the respective cell lines.

Cell Line	Growth Medium	Media Supplements	Re-seed dilution	Supplier
Human Embryonic Stem Cells	Conditioned Media	4 ng/ml basic fibroblast growth factor (bFGF)	1:3	R & D System
	mTeSR-1 ®	N/A	1:3	Stem Cell Technologies
Embryoid Bodies	Dulbecco's modified Eagle's medium (D-MEM)	20% FBS 1% L-Glutamine	N/A	Life Technologies
hESC Derived Hepatic Endoderm	Advanced RPMI 1640	B27 100 ng/ml Activin A 50 ng/ml Wnt3A	1:3	Life Technologies Peprotech R & D Systems
	Knock Out DMEM	20% Knock out Serum Replacement 0.5% L-Glutamine 1% Non-Essential Amino Acids 0.1 mM β-Mercaptoethanol 1% DMSO		all Life Technologies Sigma Aldrich
hESC Derived Hepatic Endoderm	Leibovitz L-15 Medium	8.3 % Tryptose phosphate broth 8.3% FBS 10 µM hydrocortisone 21-hemisuccinate 1 µM Insulin (bovine pancreas) 1% L-Glutamine 0.2% Ascorbic Acid 10 ng/ml Hepatocyte Growth Factor 20 ng/ml Oncostatin M		Sigma Aldrich Life Technologies Sigma Aldrich Sigma Aldrich Life Technologies Sigma Aldrich Peprotech R & D Systems
293T	Dulbecco's modified Eagle's medium (D-MEM)	10% (v/v) Foetal bovine serum (FBS) 200 mM L-Glutamine 10 mM MEM NEAA 100 mM MEM Sodium Pyruvate 1% Penicillin-Streptomycin	1:10	Life Technologies
C3A	Minimum Essential Medium Eagle (MEM)	10% FBS 0.1mM NEAA 1mM Sodium Pyruvate	1:30	Life Technologies

2.1.3 ANTIBODIES

TABLE 2.3. Antibodies used for immunofluorescence (IF) and Western blotting (WB)

Antibody	Host	Dilution for IF	Dilution for WB	Source
Alpha Feto Protein (AFP)	Mouse	1:500	1:5000	Sigma Aldrich
Albumin	Mouse	1:500	1:5000	Sigma Aldrich
Beta Actin	Mouse	1:1000	1:10000	Sigma Aldrich
Beta Tubulin III	Mouse	1:1000	1:10000	Sigma Aldrich
Cyp 3A4	Sheep	1:100	1:1000	University of Dundee*
C-Met	Rabbit	1:50	1:500	Abcam
E-Cadherin	Mouse	1:50	1:500	Santa Cruz Biotech.
HNF4 α	Rabbit	1:50	1:500	Santa Cruz Biotech.
Muscle Actin	Mouse	1:50	1:500	DAKO
Oct 4	Mouse	1:250	1:2000	Santa Cruz Biotech.
RNF4	Goat	1:50	1:500	Santa Cruz Biotech.
SENP-1	Sheep	N/A	1:1000	R & D
SNAIL	Rabbit	1:50	1:500	DAKO
SUMO-1	Sheep	N/A	1:1000	University of Dundee*
SUMO-2	Sheep	N/A	1:1000	University of Dundee*
Ubc9	Sheep	N/A	1:1000	University of Dundee*
Ubiquitin	Sheep	N/A	1:500	University of Dundee*
Vimentin	Mouse	1:200	1:2000	University of Dundee*
Anti-Rabbit 488	Donkey	1:400	N/A	Life Technologies
Anti-Mouse 488	Goat	1:400	N/A	Life Technologies
Anti-Sheep 488	Donkey	1:400	N/A	Life Technologies
Anti-Mouse HRP	Goat	N/A	1:2000	R & D
Anti-Sheep HRP	Donkey	N/A	1:2000	R & D
Anti-Rabbit HRP	Goat	N/A	1:2000	R & D
Anti-Goat HRP	Donkey	N/A	1:2000	R & D

* Kind gift of Dr R. Hay, College of Life Science, University of Dundee, Dundee, UK.

2.1.4 OLIGONUCLEOTIDES

The custom made primers were generated by MWG Eurofins Operon, Germany.

TABLE 2.4. Oligonucleotides

Gene	Primer Sequence	Temp	Cycles
HNF4ΔNT	For-CTTTGCCATATGAATGAGCGGGACCGGATCAGCACT Rev-AGGATCCTTAAGCAACTTGCCCAAAGCGGCACAGTGG	50°C 60°C	7 20
HNF4ΔCT	For-CTTTGCCATATGATGCGACTCTCCAAAACCCTC Rev-AGGATCCTTATGCTCTGCAAGGTGGGC	50°C 60°C	7 20
HNF4 PM	For- CTTCGGCATGGCCAGGATTGCCAACCTGTTGCAGG Rev-CCTGCAACAGGTTGGCAATCCTGGCCATGCCGAAG	55°C	18
AFP	For-AGAACCTGTCACAAGCTGTG Rev-GACAGCAAGCTGAGGATGTC	55 °C	30
ALB	For-CCTTTGGCACAATGAAGTGGGTAACC Rev-CAGCAGTCAGCCATTTCACCATAGG	55 °C	30
HNF4α	For-CTGCTCGGAGCCACCAAGAGATCCATG Rev-ATCATCTGCCACGTGATGCTCTGCA	55 °C	30
CYP3A4	For-CCTTACATATACACACCCTTTG Rev-GGTTGAAGAAGTCCTCCTAAGCT	50 °C	35
TO	For - GGCAGCGAAGAAGACAAATC- Rev -TCGAACAGAATCCAACTCCC	55 °C	30
TAT	For-ACTGTGTTTGGAAACCTGCC Rev-GCAGCCACTTGTCAGAATGA	55 °C	30
APOF	For-GGAAGCGATCAAACCTACCA Rev-ATCAGCCTGACAACCAGCTT	58 °C	35
CYP7A1	For- CTGCCAATCCTCTTGAGTTCC Rev- ACTCGGTAGCAGAAAGAATACATC	57 °C	35
β-ACTIN	For-TCACCACCACGGCCGAGCG Rev-TCTCCTTCTGCATCCTGTCG	58 °C	25
NANOG	For - CACTACGACCCAGGCTTCAT Rev - CTCCGCAGCTTCTTGCTTAG	60 °C	30
pLVX	For - TAGTGAACGGATCTCGACGG Rev - TTGGCGCCTACCGGTGGATG	55 °C	30
pLenti4/ BlockIT-DEST	For - GTGACTCTGGTAACTAGAGATCCC Rev - ACTACTTGAAGCACTCAAGGCAAG	55 °C	30
SUMO-1	For - TAGGATCCA TGTCTGACCAGGAGGCAA	50°C 60°C	7 20

	Rev- TCAGACAGCGGCCGCAAGCGTAGTCTGGGACGTCGTATGGGT AAA CTG TTG AAT GAC CCC CCG TTT		
SUMO-2	**For- TAGGATCCATGGCCGACGAAAAGCCCAAG** **Rev - TCAGACAGCGGCCGCAAGCGTAGTCTGGGACGTCGTATGGGT ATCAGTAGACACCTCCCGTCTGCTGTTGGAACA**	**50°C** **60°C**	**7** **20**
Ubc9	**For - AGAACGCGTGGTATGTCGGGGATCGCCCTCAGC** **Rev- GCTAGAATTCACAGATCTTCTTCAGAAATAAGTTTTTGTTCTGAG GGCGCAAACTTCTTGGCT**	**50°C** **60°C**	**7** **20**
2A	**For - CTTATGCGCGGCCGCGTGAAACAGACTTTGAATTT** **Rev - ACGCGTGAAGGGCCCTGGGTTGGACTCCACGTCTC**	**95°C**	**N/A**
HNF4 α WT	**For - CTTCGGCATGGCCAGGATTGCCAACCTGTTGCAGG** **Rev - AGGATCCTTAAGCAACTTGCCCAAAGCGGCACAGTGG**	**50°C** **60°C**	**7** **20**
Ubc9 KO	**Top - CACCGGAAGGAGGCTTGTTTAAACTCGAAAGTTTAAACAAGC CTC CTT CC** **Bottom - AAAAGGAAGGAGGCTTGTTTAAACTTTCGAGTTTAAACAAGC CTCCTTCC**	**95°C**	**N/A**
HNF4α KO	**Top - CACCGCACTCGAAGGTCAAGCTATGCGAACATAGCTTGACCTTCGAGT GC** **Bottom - AAAAGCACTCGAAGGTCAAGCTATGTTCGCATAGCTTGACCTTCGAGT GC**	**95°C**	**N/A**
HNF4 α v.1	**For – TCAGCAACGGACAGATGTCCA** **Rev- AGTGGAGCCCCCAAGCCCCAG**	**55°C**	**25**
HNF4 α v.2	**For – GTGAGTGGCCCCGACCCAGGGGAC** **Rev- AGTGGAGCCCCCAAGCCCCAG**	**55°C**	**25**
HNF4 α v.3	**For – CGTGCCAAGCCCAGGAGGGGC** **Rev- AGGAAACTGAGGCACAGACAG**	**55°C**	**30**
HNF4 α v.4	**For – GAGTGGCCCCGACCCAGGGGA** **Rev- GAGCTTATAGGGCTCAGACCC**	**55°C**	**25**
HNF4 α v.5	**For – CAACGGACAGATGTCCACCCC** **Rev- GATGGTCGGCTGGGGGATGGC**	**55°C**	**30**
HNF4 α v.6	**For – CGTGCCAAGCCCAGGAGGGGCG** **Rev- TCTCCTGGGGAGTCCCCACTC**	**55°C**	**30**

2.1.5 E. COLI HOST STRAINS

TABLE 2.5. E.coli Host Strains.

E.coli Host Strain	Usage	Source
DH5 α (Gold)	DNA amplification	Stratagene, Europe
Stabl3	DNA amplification	Life Technologies, UK
BL21	His_6 tagged protein expression	Stratagene, Europe
XL-1 Blue	Point Mutagenesis DNA expression	Stratagene, Europe

2.1.6 DNA PLASMID VECTORS

TABLE 2.6. DNA plasmid vectors.

Plasmid	Antibiotic resistance	Fusion Tag	Promoter(s)	Source
pET 15b	Ampicillin	6-His Tag	T7	Novagen
PCR 2.1	Ampicillin	N/A	T7	Life Technologies
pLVX	Ampicillin Puromycin	N/A	CMV	Clontech
pENTR™/H1/TO	Kanamycin Zeocin™	N/A	H1/TO	Life Technologies
pLenti4/BLOCK-iT™DEST	Ampicillin Zeocin™	N/A	RSV/5'LTR Hybrid	Life Technologies

2.2 MAMMALIAN CELL CULTURE AND DIFFERENTIATION

All cell culture reagents were GIBCO® products supplied by Life Technologies (UK) unless stated otherwise. Greiner (UK) supplied the plasticware utilised throughout the cell culture.

2.2.1 HUMAN EMBRYONIC STEM CELL CULTURE

Gradual Transition of hESCs to MT

H1 and H9 cells were cultured on Matrigel™ (BD Biosciences, UK) coated plasticware in conditioned media (CM, R & D Systems, UK) before being transferred into mTeSR-1® (MT, Life Tech., UK) (Table 2.2). The cells were split at a 1:3 ratio and allowed to settle overnight. They were then transferred into 80:20 ratio of CM to MT followed by 60:40, 40:60, 20:80 and finally 100% MT (Hannoun et al., 2010).

Culturing hESCs

Human ESCs were cultured for over 30 passages on Matrigel™ coated 6 well plates and were fed 4 ml of MT or CM + 4 ng/ml bFGF (Table 2.2). The media in the plates was aspirated off and 4 ml of fresh media was added. The cells were incubated at 37°C in 5% (v/v) CO_2, 95% (v/v) air, for optimal growth.

Passaging hESCs

hESCs were split at a ratio of 1:3 and were passaged using collagenase. The existing media was aspirated off, and the cells were washed once with PBS. 1ml of collagenase was added and the cells incubated at 37°C for about 3 minutes until the

edges of the colonies rounded up. The enzyme was aspirated off and the cells were washed once with PBS. 4 ml of fresh media was added to the cells and they were subsequently scraped off and triturated 1-2 times before transferring 1 ml into a new plate containing 2ml of fresh media for a 1:3 split.

Freezing and thawing hESCs

The freezing mix used consisted of knock-out serum replacement and 8% dimethyl sulfoxide (DMSO). At 90-100% confluency, hESCs were scraped off, placed into a 15ml tube, and centrifuged at 1000 rpm for 5 minutes. The supernatant was aspirated and the cells were resuspended in 0.5 ml of the freezing mix and transferred to a cryotube. The cells were stored at -80°C overnight and then transferred to liquid nitrogen.

The cells were routinely thawed by placing the cryotube in a water bath at 37°C. The cells were then taken up in 1 ml of CM and were resuspended in 6 ml of CM. Subsequently, the cells were spun at 1000 rpm for 5 minutes and resuspended in 4ml of fresh CM containing bFGF or MT and placed into a Matrigel™ coated well.

2.2.2 EMBROID BODY FORMATION

Embroid bodies (EB's) can be generated when hESCs are at about 90-100% confluent. The media was aspirated off and the cells were washed once with PBS. 4 ml of EB media was added (Table 2.2). The cells were scraped off using a cell scraper. The full 4 ml containing hESCs were placed in low cluster plates to promote cell aggregation. The EB's were fed with fresh EB media every other day for 7 days

until the EB's were very defined and vacuolated. The EB's were then transferred to 0.5% gelatin coated chamber slides (BD Biosciences, UK). The plated down EB's were allowed to differentiate spontaneously for 14 days and fed every other day with EB media. After 14 days the differentiated cells were fixed in 4% Paraformaldehyde (PFA) (Sigma Aldrich, UK) and were stained using antibodies for the three germ layers (Table 2.3), (Fletcher et al., 2008).

2.2.3 HEPATIC DIFFERENTIATION OF HESCS

hESCs were cultured and propagated on Matrigel™ coated plates with mouse embryonic fibroblast (MEF)-CM supplemented with bFGF or MT. Hepatic differentiation was initiated when hESCs reached a confluency level of approximately 30% by replacing the culture media with priming medium RPMI 1640-B27 supplemented with 100 ng/ml Activin A and 50 ng/ml Wnt3a (Table 2.2). The cells were cultured in priming medium for 3 days (the medium was replaced every 24 hours), and final priming medium with Activin A and Wnt3a was made up fresh each day. After 72 hours in priming medium, the medium was changed to differentiation medium SR-DMSO for 5 days (the medium was replaced every 48 hours). The cells were subsequently cultured in maturation and maintenance medium L-15 (Table 2.2), supplemented with 10 ng/ml hHGF and 20 ng/ml OSM for 9 days (the medium was replaced every 48 hours). Maturation and maintenance medium with hHGF and OSM was made up fresh each day. The cells gradually exhibit morphological changes from a spiky/triangular shape to a characteristic liver morphology displaying a polygonal appearance signifying hepatic endoderm (HE) formation (Hay et al., 2008).

2.2.4 THE C3A CELL LINE CULTURE

C3A cells were cultured on plastic in the presence of MEM media (Table 2.2). The cells were passaged once a week using Trypsin/EDTA at a ratio of 1:30. The C3A cells were maintained at 37°C in 5% (v/v) CO_2, 95% (v/v) air for optimal growth. The cells were cultured in T25 flasks followed by bulking up in T75 flasks. Most experiments were carried out in 6 well plates.

2.2.5 293FT CELL LINE CULTURE

293FT cells were maintained in DMEM media with supplements at 37°C in 5% (v/v) CO_2, 95% (v/v) air in T75 flasks (Table 2.2). 293FT cells were passaged using trypsin/EDTA solution and incubated for 1-5 minutes at room temperature until cells detached. Once the cells had detached, 8 ml complete medium containing Geneticin® was added and the cell suspension was transferred into a 15 ml sterile, conical tube. The cells were subsequently seeded at a ratio of 1:5 by diluting in pre-warmed complete medium containing 500µg/ml Geneticin®, and this was continuously added to the complete media throughout culture.

2.2.6 CULTURE OF FRESHLY ISOLATED PRIMARY HUMAN HEPATOCYTES

Freshly isolated primary human hepatocytes, purchased from Life Technologies (HMFS01), were maintained in Williams E media (A1217601) and the supplement pack (CM4000) on Matrigel™ (BD Biosciences, UK) coated 12-well plates. The PHHs were seeded at 105 viable cells/cm^2 and incubated at 37°C in 5% (v/v) CO_2, 95% (v/v) air.

2.3 CHARACTERISATION OF HESCS, HESC DERIVED HE AND C3A CELLS

2.3.1 IMMUNOCYTOCHEMISTRY

hESCs and hESC derived HE were washed twice with PBS, 5 minutes each wash. The HE was fixed with 4% PFA for 20 minutes at room temperature (the cells can then be stored in PBS at 4°C and stained at a later date). The cells were washed twice with PBS, 5 minutes each wash and incubated for 2 minutes at room temperature with 100% ethanol for nuclear staining. The cells were then washed twice with PBS, 5 minutes each wash. PBS/T (0.1% Tween)/10% Serum was used to block the cells for 1 hour at room temperature. The serum was removed and the respective primary antibody diluted in 1% serum (made up in PBS/T) was added and incubated for 2 hours at room temperature, or overnight at 4°C with agitation (For primary antibody details, see Table 2.3). The cells were washed 3 times with PBS at room temperature, 5 minutes each wash. The secondary antibody Alexa Flour 488 (1:400) (Table 2.3) diluted in PBS was added to the cells and incubated at room temperature for 1 hour in the dark with agitation. The cells were then washed 3 times with PBS, 5 minutes for each wash and in the dark. Each well was subsequently mounted with 50 µl MOWIOL 4-88 and DAPI (1:1000). The well was covered with a cover slip and stored at 4°C in the dark.

2.3.2 FLUORESCENCE ACTIVATED CELL SORTING

Fluorescence activated cell sorting (FACS) was used to confirm the cell surface marker expression of hESCs cultured in CM and MT. hESCs were treated with

trypsin/EDTA for 5 minutes and were lifted as single cells. Single hESCs were harvested and resuspended in FACS-PBS (PBS supplemented with 0.1% BSA and 0.1% sodium azide), counted, and resuspended at 1 x 10^{7}cells/ml for use. Aliquots of 1 x 10^{6} cells were incubated for 40 min at 4°C with the primary antibody to SSEA-4 (DHSB, Iowa; 1:200 dilution), SSEA-1 (DHSB, Iowa; 1:200 dilution), TRA 1–60, TRA-1–81 (Chemicon, Temecula, CA, both 1:200 dilution). Cells were then washed twice, removing any unbound antibody, and were then resuspended in 100 ml of FACS-PBS. Binding of primary antibody was detected using the optimum concentration (determined by titration) of an appropriate isotype specific fluorochrome labeled secondary antibody: anti-mouse IgM-PE and anti rat IgM-PE (Jackson Labs, West Grove, PA) and anti-mouse IgG3-FITC. The cells were incubated with the secondary antibody for 40 minutes at 4°C, and were washed twice and resuspended in a final volume of 250 ml. Unstained cells and cells only labeled with secondary antibody were used as controls. Dead and apoptotic cells along with debris were not included in the analysis. This was carried out by using an electronic live gate on forward scatter and side scatter parameters. Data for 5000–100,000 "live" events were acquired for each sample using a FACS Caliber cytometer equipped with a 488-nm laser and analyzed using CellQuest software (Becton Dickinson, San Jose, CA) (Fletcher et al., 2008).

2.3.3 KARYOTYPING HESCS

Chromosomal Preparations

A confluent flask of the hES cell line of interest was split and approximately 24-48hrs later Colcemid (100µl per 10ml of culture medium 100 ng/ml) was added.

After a 1.5-2.5 hour incubation the cells were harvested. The culture media was initially transferred into a 15 ml conical bottom centrifuge tube. The cells were washed with 3 ml of PBS and the solution was added into the tube containing the initial media. 1 ml of trypsin solution was added to the flask and incubated at 37°C until the cells lift as single cells. The culture medium in the centrifuge tube was used to wash the cells in trypsin off the flask and the resulting cell suspension was collected in the centrifuge tube and centrifuged at 1300 rpm for 5 minutes. The supernatant was decanted into a waste beaker and the cell pellet was resuspended in 10ml of 0.56% KCl (hypotonic solution) using a vortex mixer. The resuspended cells were incubated at room temperature for 10 minutes and then centrifuged at 1300 rpm for 5 minutes. The cell pellet was resuspended in 10 ml of fixative (3:1 Methanol: Acetic Acid) by adding the fixative slowly down the side of the tube and then centrifuged at 1300 rpm for 5 minutes. 7 ml of fresh fixative was added whilst mixing on the vortex mixer and this step was repeated once more with a final centrifugation at 1300 rpm for 5 minutes. The cells were resuspended in a small volume of fresh fixative, and the resulting suspension should be just cloudy. A pre-soaked slide was polished (slides were stored in Ethanol + 5% Hydrochloric acid), and one drop of the cell suspension was dropped, using a fine tipped pastette, onto the surface of the slide, whereby the drop should spread evenly over the slide surface. The slides were left on the bench to air dry. Microscopy analysis using the X10 phase contrast objective was used to check for the mitotic index, spreading and fixation. Once the preparation was validated, 3 additional slides were made. The remaining cell suspension was stored at –20°C for later use.

GIEMSA Staining of the Chromosomal Preparations

The slides of chromosome preparations were allowed to age at room temperature for 3-5 days. A coplin jar was filled with 2X SSC (0.3 M NaCl and 0.035 M Tri Sodium Citrate) and was incubated at 60°C for 20-30 minutes to allow the solution to warm up. The slides were then incubated in the 2X SSC at 60°C for 2-4 hours. The slides were individually washed with running tap water and incubated in 1% trypsin for 20-30 seconds. The slides were washed once again and incubated in 5% GIEMSA (40ml GURRS buffer pH6.8 and 2 ml of GURRS R66 GIEMSA stain) for 8-10 minutes with a final wash in water. The slides were blotted dry and soaked in xylene for a few hours. DePeX was used to mount the slides once they were dry and they were left overnight to allow the coverslip to set completely. The slides were stored in the dark until analysed.

2.3.4 ELISA ASSAYS

L-15 medium (1ml) was added to the cells on days 14, 15 and 16 of the differentiation protocol and incubated for 24 hours at 37°C in 5% (v/v) CO_2, 95% (v/v) air (n=3). The supernatants were collected after 24 hours and could be stored at -80°C for later use. High binding EIA plates were coated with rabbit anti-human antibodies specific for the protein overnight at 4°C (usual serum protein screen consists of fibrinogen (1:10000), fibronectin (1:1000) and thyroxin binding pre albumin (1:1000)). The sample supernatants were diluted 1:10 and pipetted into a 96 well plate in triplicate followed by a 2-hour incubation at room temperature. Peroxidase conjugated rabbit anti-human antibody specific for the appropriate

protein was added to the wells and incubated at room temperature for 1 hour. O-phenylenediamine was added into the wells for 5-15 minutes and the reaction was stopped by adding 0.5 M sulphuric acid. The plates were read at 490 nm with a reference wavelength of 630 nm using an MRX II plate reader. The data was then normalised to per mg protein as determined by the BCA Assay (Pierce, UK) (Hay et al., 2008).

2.3.5 CYTOCHROME P450 ASSAYS

Day 17 hESC (or another specified time point) derived HE and C3A cells were incubated with the luciferin conjugated specific Cyp 3A4 (1:40) and 1A2 (1:50) substrate (P450 P-Glo™ Luminescent Kit, Promega, UK) for 5 hours at 37°C (n=3). The tissue culture media was used as a negative control. The supernatants were then collected and could be stored at -80°C for later use. The Luciferin detection reagent was reconstituted by mixing the buffer into the bottle containing the lyophilized Luciferin detection reagent. 50 µl of the supernatant sample was mixed with 50 µl of the detection reagent in a white 96 well plate and incubated at room temperature in the dark for 20 minutes. The relative levels of basal activity were measured using a luminometer and were normalised to per mg protein as determined by the BCA Assay.

2.3.6 UREAGENENSIS ASSAY

hESC derived HE and the C3A cells at a defined time point were washed with 1 ml of PBS. 1 ml of PBS was added to each well and a final concentration of 16 µM of ammonium chloride solution was added to the tested wells. A PBS control was used.

The solutions were incubated for 4 hours at 37°C and the supernatants were subsequently collected and could be stored at -20°C for later use. 100 µl of the sample was added to a cuvette with 1 ml of buffer (150 ml PBS, 300 µl 0.1 M ADP, 50 mg NADH and 750 µl of 0.5 M alpha keto glutarate pH 7.4) and read at 340 nm on a spectrophotometer. Glutamate dehydrogenase was diluted 1:5 with water and 10 µl of the solution was added into the cuvette and mixed by inverting. The reaction was incubated at room temperature for 2 hours and the absorbance at 340 nm was read. Urease was diluted 1:3.5 with water and 10 µl was added into the cuvette. The reaction was inverted to allow mixing and incubated at room temperature for 2 hours. The absorbance was then measured as 340 nm. The urea concentration was calculated using the following formula: Urea (mmol/l)= Abs/6.22/2 (originated from the following formula Abs = ε x l x c where ε is the molar extinction coefficient which in this case is 6.22 $mmol^{-1}$ cm^{-1} for NADH, l is the length of the cuvette and c is the concentration). The values were normalised to per mg of protein determined by the BCA assay.

2.3.7 CASPASE 3/7 ASSAYS

Apoptosis was measured using the Caspase 3/7 –Glo-Kit (Promega, UK). The caspase 3/7 buffer was combined with the lyophilized substrate and the solution was allowed to equilibrate to room temperature. 500 µl of the caspase 3/7 reagent was added to the cells containing 500 µl of their respective media and incubated for 1 hour. The supernatants were collected and luminescence was measured using a luminometer. The measured readings were then normalised to per mg of protein determined by the BCA assay.

2.4 MOLECULAR TECHNIQUES

2.4.1 RNA ISOLATION AND EXTRACTION

The cells of interest were washed with PBS and 1 ml of TRIZOL reagent was added per well of a 6 well plate and left to incubate at room temperature for 5 minutes. The cells were collected and placed in a 1.5 ml eppendorf (stored at -80°C for later use if required). 0.5 ml of Chloroform was added to the eppendorf and mixed by inverting; this is done in a fume hood. The solution was centrifuged at 13,000 rpm for 15 minutes at 4°C. The aqueous layer was collected and placed into a clean eppendorf. 1 ml of isopropanol was added into each tube, mixed by inverting the tube and incubated at room temperature for 10 minutes to precipitate the RNA. The solution was centrifuged at 13,000 rpm for 10 minutes at 4°C. The supernatant was aspirated without disturbing the RNA pellet. 0.5 ml of 70% ethanol was used to wash the RNA. Following a 5 minute incubation at room temperature the RNA was centrifuged at 8,000 rpm for 5 minutes at 4°C. The ethanol was aspirated and the pellet was left to dry at room temperature for 5-10 minutes. Once all the ethanol had evaporated, the pellet was resuspended in 30 µl of deionised water and stored at -80°C for later use. All RNA and DNA sequences were quantified using the Nanodrop, concentration and purity were measured.

2.4.2 REVERSE TRANSCRIPTION AND POLYMERASE CHAIN REACTION (PCR)

Reverse transcription (RT) was carried out using the Promega Reverse Transcription AMV kit (Promega, UK). The purified RNA was converted to cDNA using the

reverse transcriptase enzyme AMV, as per the manufacturer's instructions. A standard RT reaction was as follows:

RNA template	200ng
Random Hexamers	1 µl
Nucleotide mix (10 mM of each dNTP)	5 µl
AMV Reverse Transcriptase	1 µl
RNAsin	5 µl
5x AMV-RT Buffer	5 µl
Nuclease Free Water	Up to a final volume of 25 µl

The resulting cDNA was used for further analysis carried out by PCR. PCR was employed using the Gold Ampli-Taq polymerase (Applied Biosciences, UK). The PCR reactions were set up using specific primers and cycle conditions specified in Table 2.4 in conjunction with the manufacturer's instructions. PCR was used to investigate the expression of specific genes within cells, this is defined as RT-PCR and utilises cDNA as the template. PCR was also employed to generate full gene sequences for vector cloning and subsequent protein expression, in this situation the DNA template was a gene specific image clone (Source Bioscience, UK). A standard PCR reaction was as follows:

DNA template (50 ng/ml)	2µl
Nucleotide mix (10 mM of each dNTP)	1 µl
Oligonucleotide 1 (10 pmol/µl)	1 µl
Oligonucleotide 2 (10 pmol/µl)	1 µl
Gold Ampli-Taq Polymerase	0.23 µl
Ampli-Taq 10x Buffer	5 µl
$MgCl_2$	Up to a final volume of 50 µl
Nuclease Free Water	

The polymerase enzyme mix amplifies the desired sequence by extending the annealed oligonucleotides to synthesise DNA from 3' end of the oligonucleotide.

This process of denaturation, annealing and synthesis is repeated a number of times, whereby each newly synthesised DNA acts as a template for the next reaction. Agarose gel electrophoresis (section 2.4.3) was used to ensure that a PCR product was visible indicative of specific primers and a successful PCR reaction.

2.4.3 GEL ELECTROPHORESIS

Agarose gel electrophoresis was employed to visualise the amplified (PCR) and manipulated DNA (digests or ligations). A 1% (w/v) agarose gel solution in 1 x Tris base/Acetic acid/EDTA (TAE, Life Technologies, UK) buffer was prepared and the solution brought to the boil to dissolve the agarose. On cooling, SYBR Safe (Life Technologies, UK) was added at a dilution of 1:10,000 and the agarose solution was poured into the gel rack containing the appropriate gel comb. Blue gel loading dye (New England Biolabs, UK) was diluted 1:6 and subsequently added to the DNA sample. The DNA samples were loaded into the gel, with a 1 Kb and 100 bp plus DNA markers (Life, Technologies, UK). DNA gels were run at 100 V in 1 x TAE buffer, and imaged using a UV light source.

2.4.4 RESTRICTION DIGESTS

Restriction endonucleases, supplied by New England Biolabs, were used to cleave specific restriction sites in plasmid vectors. This method was used for directional cloning, verifying the correct insert within a plasmid vector and for isolating the DNA sequence of interest. Double digests were only used when the enzyme buffers displayed sufficient compatibility. The digestion reactions were carried out for 2

hours at 37°C and were subsequently heat inactivated when possible. The digestion reactions consist of:

Plasmid DNA	1 µg plasmid DNA
Restriction enzyme 1	0.5 µl (enzymes 1/10 final volume)
Restriction enzyme 2	0.5 µl
10 X Enzyme Buffer	1 µl
BSA (10 mg/ml)*	1 µl
Sterile double distilled H_2O	Up to a final volume of 10 µl

*BSA was only added when required, as suggested by the manufacturer's instructions.

The digestion reaction efficiency was validated by gel electrophoresis (section 2.4.3). The DNA band was excised and gel purified using the QIAGEN gel extraction kit (section 2.4.5), (QIAGEN, UK) when the insert was required for subsequent cloning.

2.4.5 DNA GEL EXTRACTION

The QIAGEN gel extraction kit (QIAGEN, UK) was implemented for DNA purification from an agarose gel. The DNA band of interest was carefully excised using a scalpel and extracted as instructed by the manufacturer.

2.4.6 LIGATION REACTIONS

T4 DNA ligase is the ideal enzyme required for the efficient ligation of 'sticky' ended sequences, usually an insert of interest and a plasmid vector. The most optimum ligation ratio was found to be 1:3 (vector: insert), unless otherwise stated. The ligation reaction was incubated at room temperature for 1 hour, however in

special cases ligation was carried out at 16°C overnight. The ligation reaction was set up in accordance with the manufacturer's instructions and consisted of the following reagents:

Vector DNA	1 µl
Insert DNA	3 µl
T4 Ligation Buffer (2X)	5 µl
T4 DNA Ligase	1 µl
Nuclease Free Water	Up to a final volume of 10 µl

Successfully ligated inserts were verified using gel electrophoresis (section 2.4.3), and were subsequently transformed into E.coli bacterial cells.

2.4.7 TRANSFORMATION REACTIONS OF PLASMID DNA

E.coli strains have been optimized to propagate plasmid DNA and express proteins of interest, each strain has unique properties required for effective function. The bacterial strains used are mentioned in Table 2.5. 10 µl of the ligation reaction was added to 50 µl of the competent cells at 4°C and was incubated for 10 minutes. The cells were then heat shocked in a 42°C water bath for 45 seconds followed by a 2 minute incubation on ice. 0.25 ml of super optimal broth (SOC) media was added into the eppendorf and the bacterial cells were cultured at 37°C for 1 hour. Following bacterial propagation, the cell suspension was plated in 100 µl and 50 µl volumes on agar plates containing 0.1 mg/ml of the specific antibiotic. The plates were incubated overnight at 37°C. Selection pressure ensures that the bacterial cells retain the plasmid of interest as it contains the antibiotic resistance gene required for survival.

2.4.8 SITE DIRECTED POINT MUTAGENESIS

The QuikChange II Site-Directed Mutagenesis Kit (Stratagene, UK) was utilized to make defined point mutations in vitro to cloned DNA, as noted by the manufacturer's instructions. In this specific case, an extension time of 8 minutes was used (1 min per kb) and 18 cycles defined the PCR parameters. The reaction consisted of the following reagents:

10X Reaction Buffer	5 µl
DNA template	5 ng/µl of plasmid DNA
Oligonucleotide 1	100 ng/µl
Oligonucleotide 2	100 ng/µl
Nucleotide mix (10 mM of each dNTP)	1 µl
Pfu Ultra DNA polymerase	2.5 U/µl
Nuclease Free Water	Up to a final volume of 50 µl

Agarose gel electrophoresis (section 2.4.3) verified the generation of a PCR product. *Dpn*I (1 µl) restriction enzyme was added directly to each amplification reaction and was incubated at 37°C for 1 hour to digest the parental supercoiled dsDNA. The PCR product was subsequently transformed into XL-1 Blue super-competent *E. coli* cells (section 2.4.7). Blue white selection was employed, whereby the control pWhitescript™ plasmid was used to confirm successful mutagenesis, indicated by the blue colonies. The XL-1 Blue cells containing the mutagenesis control were plated on agar plates containing ampicillin selection in addition to IPTG and X-galactocidase.

2.4.9 PLASMID DNA PURIFICATION

Small Scale

The QIAprep Miniprep Kit (QIAGEN, UK) was used to isolate and purify plasmid DNA in a small scale according to the manufacturer's instructions. A single colony of

transformed *E.coli* containing the DNA plasmid of interest was picked and used to inoculate a culture of 5 ml of LB medium, containing the appropriate antibiotic. The culture was incubated overnight at 37ºC with shaking at 200 rpm, allowing adequate propagation of the plasmid vector.

The bacterial cells were harvested in 15 ml tubes by centrifugation at 5000 x*g* in a swinging bucket conventional table top centrifuge for 5 min at 4 ºC. The pelleted bacterial cells were re-suspended in 250 µl of Buffer P1 and transferred to a micro-centrifuge tube. 250 µl of Buffer P2 was added and the tube inverted 4-6 times, followed by the addition of 350 µl of Buffer N3 and again the tube was inverted 4-6 times. The solution was centrifuged at 13,000 x*g* in a micro-centrifuge for 10 min at room temperature. The supernatant was applied to a QIAprep spin column by pipetting and centrifuged for 30-60 sec, discarding the flow through. The QIA prep spin column was washed by the addition of 0.5 ml of Buffer PB and centrifuged for 30-60 sec, and again the flow through was discarded. The QIAprep spin column was further washed with 0.75 ml of Buffer PE and centrifuged for 30-60 sec, and the flow through discarded. To elute the DNA, the QIAprep column was placed in a clean microcentrifuge tube and 30 µl of Buffer EB added to the centre of the column. The column was allowed to stand for 1 min, followed by centrifugation for 1 min.

Large Scale

A single colony was picked from a freshly streaked selective plate and a starter culture of 2–5 ml LB medium containing the appropriate selective antibiotic was inoculated and incubated for 8 hours at 37°C with vigorous shaking (approx. 300

rpm). The starter culture was diluted 1:500 to 1:1000 into selective LB medium and were cultured at 37°C for 12–16 hours with vigorous shaking (approx. 300 rpm).

The bacterial cells were harvested by centrifugation at 6000 x *g* for 15 min at 4°C, resuspended in 10 ml of Buffer P1 containing RNase A and inverted 4-6 times. 10 ml of Buffer P2 was added and the solutions were mixed thoroughly by vigorously inverting the sealed tube 4–6 times. The solutions were incubated at room temperature for 5 minutes to ensure efficient lysis. During the incubation period, the QIAfilter Cartridge was prepared by screwing the cap onto the outlet nozzle of the QIAfilter Maxi Cartridge. 10 ml of chilled Buffer P3 was added to the lysate, and mixed immediately by vigorously inverting the tube 4–6 times. The lysate was poured into the barrel of the QIAfilter Cartridge and incubated at room temperature for 10 minutes. A HiSpeed Maxi Tip was equilibrated by applying 10 ml Buffer QBT and the column was allowed to empty by gravity flow. The cap from the QIAfilter outlet nozzle was removed and the plunger was gently inserted into the QIAfilter Maxi Cartridge. The cell lysate was filtered into the previously equilibrated HiSpeed Tip and allowed to enter the resin by gravity flow. A HiSpeed Maxi Tip was washed with 60 ml Buffer QC. The DNA was eluted with 15 ml of Buffer QF and collected in a 50 ml tube. The DNA was precipitated by adding 10.5 ml (0.7 volumes) of room temperature isopropanol. The solution was mixed and incubated at room temperature for 5 minutes. During incubation the plunger was removed from a 30 ml syringe and attached onto the outer nozzle of the QIAprecipitator Maxi Module. The QIAprecipitator was placed over a waste bottle and the eluate/isopropanol mixture was transferred into the 30 ml syringe. The plunger was then inserted and the mixture

was filtered through the QIAprecipitator using constant pressure. The QIAprecipitator was removed from the syringe and the plunger was pulled out. The QIAprecipitator was re-attached and 2 ml of 70% ethanol was added to the syringe. The DNA was washed by inserting the plunger and pressing the ethanol through the QIAprecipitator using constant pressure. The QIAprecipitator was once again removed from the syringe and the plunger was pulled out. The QIAprecipitator was attached to the syringe again, the plunger was re-inserted, and the membrane was dried by pressing air through the QIAprecipitator quickly and forcefully. This step was repeated. The outlet nozzle of the QIAprecipitator was dried with absorbent paper to prevent ethanol carryover. The plunger was removed from a new 5 ml syringe and attached onto the outlet nozzle of the QIAprecipitator. The outlet of the QIAprecipitator was held over a 1.5 ml collection tube and 0.5 ml of Buffer TE was added to the 5 ml syringe. The plunger was inserted and the DNA was eluted into the collection tube using constant pressure. The QIAprecipitator was removed from the 5 ml syringe, the plunger was pulled out and the QIAprecipitator was re-attached to the 5 ml syringe. The initial eluted solution was transferred into the syringe for a second elution to ensure high yields of the DNA.

2.4.10 DNA SEQUENCING

MWG Eurofins Operon, Germany, carried out the sequencing reactions. Sequencing utilized specified primers provided by the company or custom designed primers when unavailable. The DNA sequences were analysed with Invitrogen Vector NTI software using the basic alignment tool.

2.4.11 SOLEXA

SOLEXA is a deep sequencing technique used to measure the absolute levels of transcript expression within cells. The cDNA obtained from the purified RNA is initially cleaved to a uniform length between 200 and 500 bases and then tagged using unique adapters which then bind to the flow cell in clusters. The flow cell consists of 8 lanes whereby each set of samples can be run simultaneously. The cDNA is then sequenced, whereby each cycle adds a nucleotide to the sequence. A fixed camera with laser capabilities reads the nucleotide using illumination. The base addition cycle is repeated after de-blocking and fluorophore removal from the terminal base. The final sequence is a result of aligning 18-50 sets of reads. Paired end reads were carried out in this run, generating 50 base reads from both ends of the template. Pair end sequencing is useful for differentiating gene splice variants within a sample. The data is then analysed using specific software programs and normalised to allow definitive comparisons of the genes between the six samples. Two types for normalisation were used. Firstly, the raw count of reads for each gene was divided by the total number of reads aligned and multiplied by one million, thus accounting for the depth of sequencing. Secondly, the results were divided by the length of the gene multiplied by one thousand to obtain a read per kilo bases per million reads value, RPKM value.

2.5 PROTEIN BIOCHEMISTRY TECHNIQUES

2.5.1 CELLULAR PROTEIN EXTRACTION

Cells grown in a 6 well plate were lysed in 150 µl of SUMO lysis buffer (2% SDS (Sigma Aldrich), 50 mM Tris pH 8 (Sigma Aldrich), 1 mM ethylenediaminetetraacetic acid (EDTA, Sigma Aldrich) and 10 mM iodoacetamide

(Sigma Aldrich)) for 5 minutes at room temperature. The cell extracts were sonicated and stored at -80°C for later use.

2.5.2 MEASURING PROTEIN CONCENTRATION

The Pierce BCA (bicinchoninic acid) protein assay was used to quantify the protein concentration in the cell protein extract samples. Protein extracts were diluted 1:5 using nuclease free water (2 µl of sample extract and 8 µl of water) in a 96 well plate, each sample was pipetted in duplicate. Reagents A and B were mixed at a 50:1 ratio and a volume of 200 µl was transferred into each sample well in addition to wells containing the bovine serum albumin standards ranging from 20-2000 µg/ml, as per the manufacturer's instructions. The plate was incubated at 37°C for 10 minutes and the absorbance was read at 562nm. The protein concentrations were calculated by linear extrapolation using the standard curve generated from the protein standards.

2.5.3 SDS-NUPAGE® POLYACRYLAMIDE GEL ELECTROPHORESIS

The SDS-NuPage® gel electrophoresis (SDS-PAGE) was used to separate proteins of varying molecular weights. The XCell SureLock® Mini-Cell system (Life Technologies, UK) was used with 4-12% Bis-Tris pre-cast polyacrylamide gels (Life Technologies, UK). The samples were denatured at 70 ºC for 10 minutes in 4 x LDS sample buffer (Table 2.1). Once the gel was fitted in the chamber, the tank was filled with 1x NuPage® MES-SDS running buffer in addition to 0.5 ml of antioxidant (Life Technologies, UK) in the inner chamber. The samples were loaded, including kaleidoscope® Precision Plus Protein Standards marker (Biorad, UK). A current of 200 V was applied and the samples were run for approximately 1 hour. The gels

containing the separated proteins were carefully removed from the cassette and were either Commassie stained (section 2.5.5), or used in western blotting (section 2.5.4).

2.5.4 WESTERN IMMUNOBLOTTING

Protein Transfer

Western blotting was used to detect the presence of specific proteins in samples extracts using the XCell SureLock® Mini-Cell system, as per the manufacturer's instructions. Proteins were separated via SDS-PAGE followed by subsequent transfer from the polyacrylamide gel to the Polyvinylidene fluoride (PVDF) membrane, prevents protein cross over if the transfer time was over run and provides long-term durability (Millipore, UK). The membrane was then probed with antibodies specific to the target protein. The transfer stack was assembled in the following order from cathode to anode: 2x sponge; filter paper soaked in 1x transfer buffer; SDS-PAGE gel; PVDF membrane pre-soaked in methanol and then transfer buffer on top of the gel; filter paper soaked in transfer buffer; 3x sponge, and the beginning sequence was repeated for a second gel. It is vital that the membrane is positioned accurately between the gel and the anode as the samples and current will be moving in that direction. The stack was assembled in the XCell Blot II module and was tightly sealed and placed into the transfer SureLock® tank containing 1x transfer buffer (Table 2.1) and 0.25 ml of antioxidant in the inner chamber and cold water in the outer chamber. A constant current of 160 mA was applied for 90 min.

Immunoblotting

Once the proteins have been successfully transferred onto the PVDF membrane, the membrane was blocked to prevent non specific antibody binding. The membrane was blocked in 30 ml of 5-10% (w/v) dried milk, depending on the specificity of the antibody, in 0.1% (v/v) PBS Tween (PBST) for 1 hour with gentle agitation at room temperature. The primary antibody was added to 3 ml of the milk + PBST at the appropriate dilution for the antibody (Table 2.3), and incubated overnight at 4°C with gentle agitation. Unbound antibody was removed by three, 5-minute washes with 50 ml of PBST. A horseradish peroxidase (HRP) -conjugated secondary antibody was diluted in 10 ml of 5% dried milk in 0.1% PBST at a dilution of 1:2000, and incubated for 1 hour with gentle agitation at room temperature. Unbound antibody was, once again, removed by three, 5-minute washes with 50 ml PBST.

Enhanced Chemiluminescence (ECL)

The proteins of interest were detected using enhanced chemiluminescence (ECL). The HRP substrate reacts with the conjugated HRP group present on the secondary antibody specifying the target protein, whereby x-ray films detect the signal released. Protein bands were visualised using the Peirce enhanced chemiluminescence reagent kit (Pierce, UK). Peroxide Buffer and the Luminol/Enhancer Solution was mixed at a 1:1 ratio and spotted on to the membrane (2 ml for each membrane), ensuring it was evenly spread. In the dark room, the membrane was exposed to Kodak chemiluminescence BioMax MS x-ray film for the appropriate length of time, typically 1-30 minutes. The exposed film was developed using a film developer containing the fixative, developer and water solutions.

2.5.5 COMMASSIE GEL STAINING

Proteins were visualised by Coomassie Brilliant Blue R-250 staining after separation using SDS-PAGE electrophoresis. The SDS-PAGE gel was placed in a glass container containing 5 ml of Commassie stain (Table 2.1). The gel was incubated with stain for approximately 1-2 hours with gentle agitation at room temperature. Longer staining periods were used if more pronounced bands were required. The gels were washed with water and placed in destain solution (Table 2.1) until all background staining was removed. The gels were once again washed with water and pictures were taken to record the protein patterns.

2.5.6 CROSS LINKING ANTIBODIES

Cross-linking antibodies to beads allows specific pull down of the protein of interest leaving behind all non-specific bound proteins. The HNF4α antibody was cross linked to protein A beads (Generon) using the protocol stated in the book 'Using Antibodies: a laboratory manual' by Ed Harlow, 1999. The antibodies and beads were incubated for 1 hour at room temperature. The beads were washed with 0.2 M sodium borate solution and were further incubated with dimethyl pimelimidate for 30 minutes at room temperature to allow cross links to form. Ethanolamine was used to stop the reaction and the beads were washed with 0.01% Merthiolate PBS. Coupling efficiency was verified using SDS-PAGE and Commassie gel staining to confirm the disappearance of the antibody heavy chain band (55 kDa).

2.5.7 IMMUNOPRECIPITATION

The cell extract was diluted using buffer (50 mM Tris pH 8, 150 mM NaCl, 5 mM EDTA and 2 mM DTT, and 1% Nonidet-P40 (NP40), all Sigma Aldrich). The cell extract was incubated with the 1.5 µg of the HNF4 α antibody (Santa Cruz) and a rabbit IgG (Santa Cruz) was used as a control, for 3 hours at 4°C. Protein A beads (Sigma Aldrich) were prepared in NP40 buffer (50mM Tris pH8, 150mM NaCl, 5mM EDTA and 1%NP40, all Sigma Aldrich). 20 µl of beads were added to each sample and left overnight at 4°C under constant rotation. The samples were spun down at 3500 rpm for 2 minutes, washed 3 times with NP40 buffer and eluted with 2X boiling mix at 100°C for 2 minutes. The samples were then spun at 13,000 rpm for 5 minutes and analysed by Western Blot (SUMO 2 antibody).

2.5.8 PROTEIN EXPRESSION

HNF4 α wild type protein was kindly provided to us by Primorigen Inc. C and N terminal deletions and the point mutant sequences of the HNF4 α were generated by amplifying the respective fragment using primer specific PCR; please refer to Table 2.4 for primer details. A two cycle PCR program was used where initially the specific sequence was generated to be the predominant template in the reaction (50°C for 7 cycles) followed by amplification of the sequence (60°C for 20 cycles) (section 2.4.1). The C terminal deletion lacked the last 115 amino acids whilst the N terminal deletion lacked the first 140 amino acids, as did the point mutant. The fragments were cloned into the pET15b Vector (Novagen, UK), using specific Bam HI and Nde I restriction sites (section 2.4.4-2.4.6). All mutants contained a His tag to confirm the expression and aid in the purification process. The constructs were then

transformed into BL21 E-coli (Stratagene, UK) using chemical induced transformation (section 2.4.7). The BL21 cells were grown at 37˚C until they reached an optical density of 0.5 and were induced with 0.5 mM IPTG (Life Technologies, UK) and left for 1 hour. The cells were spun down and were lysed in SUMO lysis buffer. The proteins were then purified as previously described by Hoffman et al. (Hoffmann et al., 1991) using specific nickel-agarose beads for His tagged protein pull down (section 2.5.9). Point Mutagenesis was carried out as directed by the Quick Change II manual (section 2.4.8) (Stratagene, UK), and the protein was expressed and purified as described above.

2.5.9 PROTEIN PURIFICATION

All protein purification steps were carried out at 4^{o}C. The bacterial cells were harvested by centrifugation and resuspended in 30 ml of low salt buffer (Table 2.1), supplemented with protease inhibitors. N-P40 was added to a final concentration of 1% and the cell suspensions were mixed thoroughly and incubated for 3 minutes to ensure good cell lysis. The lysed cells were sonicated on ice repeatedly, followed by centrifugation at 16,000 *g* for 30 min. The supernatant was carefully transferred into a fresh falcon tube and maintained on ice. The remaining pellet was fast frozen on dry ice and stored at -80^{o}C. Immidazole was added to the cleared lysate to a final concentration of 5 mM. The solution was then applied to pre-equilibrated nickel agarose (1 ml of resin) and incubated for 1 hour at 4^{o}C with constant rotation. The resin was transferred into a fresh column and was allowed to pack down using gravity. The bound protein was washed with 50 bed volumes of chilled CB100 followed by 50 bed volumes of chilled CB1000 (Table 2.1). The protein was eluted

using 10 ml of chilled elution buffer (Table 2.1) and 0.5 ml fractions were collected in 15-20 eppendorf tubes. 10 µl of each fraction was transferred into a well of a 96 well plate with the elution buffer used as a negative control. 90 µl of Bradford reagent (Life Technologies, UK) was added into each well. The fractions containing the eluted protein turned blue. The fractions were then pooled together and dialyzed overnight at 4^{o}C in dialysis buffer (Table 2.1), with constant stirring. The remaining purified protein was aliquoted and stored at -80 ^{o}C for later use.

2.6 IMMUNOHISTOCHEMISTRY – SECTION STAINING

IHC was used to stain for specific proteins expressed in adult liver sections. DAB staining was preferred over fluorescence, as fluorescence results in high background signal due to non-specific binding of the secondary antibody. 3,3'-Diaminobenzidine (DAB) is an organic compound used as a positive indicator in immunohistochemistry. The secondary antibody is conjugated to a peroxidase enzyme which binds DAB as a substrate and results in its oxidization and a subsequent colour change to dark brown.

Antigen Retrieval

The wax sections on glass slides were placed in a slide holder and incubated in xylene for 10 minutes with intermittent agitation to allow adequate removal of the wax. The xylene solution was blotted off using paper towels and the slides were placed in 100% ethanol for 2 minutes with intermittent agitation. The slides were transferred into ethanol solutions with decreasing concentration; 90%, 80%, 70% and 50%, and each incubation was 2 minutes with intermittent agitation, this resulted in

section re-hydration. The slides were finally placed in water, and it is essential that the slides never dry.

Both enzymatic and heat antigen retrieval techniques were employed, however, with the RNF4 antibody in use, the trypsin based antigen retrieval method was found to be the most optimal. 1% Trypsin solution was made up in water containing 1% $CaCl_2$ and the slides were incubated for 15 minutes at room temperature.

DAB Staining

The slides were then blocked with 120 µl of hydrogen peroxidase blocking solution for 15 minutes at room temperature, this steps prevents any false positive signals. The slides were washed twice with PBS and 120 µl of blocking protein serum was added for 20 minutes at room temperature. The primary antibody was added at 1:20, diluted in DAKO serum diluent, and was incubated overnight at 4°C. After overnight incubation the slides were washed twice with PBS and 3 drops of the HRP anti Rabbit secondary antibody solution (Invision, UK) were added to the slides and incubated for 1 hour at room temperature. The slides were washed twice with PBS and DAB was placed onto the slide and left for 3-5 minutes at room temperature. The slides were counter-stained with hematoxin, which stains nuclei purple, for 25 seconds at room temperature. The slides were dehydrated by incubation in increasing concentrations of ethanol solution followed by a 5-minute incubation in 100% ethanol at room temperature followed by another 5 minute incubation in xylene. DPX mountant was dropped onto the slides and covered with a coverslip and left to dry overnight. The sections were then observed using bright field microscopy.

2.7 IN VITRO PROTEIN ASSAYS

2.7.1 IN VITRO SUMOYLATION ASSAY

An in vitro SUMOylation assay was used to verify the ability of a protein to be SUMO modified in the presence of the SUMO conjugation machinery. 3 µg of each HNF4 α variant and the Ran GAP positive control was conjugated in a reaction containing: 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM DTT, 5 mM MgCl2, 2 mM ATP, 100 ng SAE2/SAE1, 400 ng Ubc9 and 2.5 µg SUMO-1 or SUMO-2 and incubated at 37°C for 3-4 hours. SUMOylation was verified using western blotting (section 2.5.3 and 2.5.4) (Tatham et al., 2008).

2.7.2 IN VITRO SUMO DECONJUGATION ASSAY

Deconjugation assays were employed as a further control to confirm sufficient SUMO conjugation. SENP enzymes were utilized to cleave any isopeptide bonds between SUMO and the target proteins. 10 mM of iodoacetamide was added to the completed conjugation reaction and left at room temperature for 30 minutes. DTT was added at 20 mM to the reaction and left for 15 minutes at room temperature. 10 nM-50 nM of the respective SENP enzyme was added into the reaction and incubated for 1 hour at 37°C. Deconjugation was validated using western blotting (section 2.5.3 and 2.5.4) (Shen et al., 2009).

2.7.3 IN VITRO UBIQUITINATION ASSAY

1 µg of HNF4 α was either mono or poly-SUMOylated as described above. The ubiquitination machinery (8 µM of ubiquitin, 40 nM Uba1, 0.7 µM UbcH5a, 0.5 µM RNF4) was added to the reaction to make a total volume of 50 µl under the following

conditions: 50 mM Tris, 200 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 2 mM ATP and 0.1% NP40. The reaction was incubated at 37°C for 3-4 hours. The HNF4α was immunoprecipitated (section 2.5.7), out of the reaction using the cross-linked beads specific to HNF4α and the elute was analysed using western blotting (section 2.5.3 and 2.5.4).

2.8 LENTIVIRAL BASED TECHNIQUES

2.8.1 LENTIVIRAL KNOCK DOWN VECTOR CONSTRUCTION

The BLOCK-iT™ Inducible Lentiviral RNAi System (Life Technologies, UK) was used to specifically knock down genes expressed in vivo. The vectors were generated as described by the manufacturer's instructions. Specific shRNA sequences were designed (Table 2.4), and annealed at 95°C for 4 minutes in the following reaction:

200 µM top strand oligonucleotide	5 µl
200 µM bottom strand oligonucleotide	5 µl
10X Annealing Buffer	2 µl
Nuclease Free Water	8 µl

The annealed sequences were ligated into the pENTR™/HI/TO entry vector in the reaction set up below, and the reaction was incubated for 1 hour at room temperature.

5X Ligation Buffer	4 µl
pENTR™/H1/TO (0.75 ng/µl)	2 µl
Annealed reaction (5 nM; 1:10,000 dilution)	1-5 µl
T4 DNA Ligase (1 U/µl)	1 µl
Nuclease Free Water	to a final volume of 19 µl

The ligation reaction was transformed into DH5α cells as previously described in section 2.4.7, and the vector containing the insert was purified as mentioned in

section 2.4.9. The ligated vectors containing the shRNA were verified by sequencing (section 2.4.10).

The entry vector provides an efficient platform for inserting the shRNA sequences into the pLenti4/BLOCK-iT™-DEST vector. This was done by a recombination reaction set up as below:

Entry clone (50–150 ng/reaction)	1-7µl
pLenti4/BLOCK-iT™-DEST vector (150 ng/µL)	1µl
TE Buffer, pH 8.0	Up to 8µl
Gateway® LR Clonase II Enzyme Mix	2µl

The reaction was incubated at room temperature for 1 hour and subsequent stopped by the addition of 1 µl of Proteinase K solution, which was incubated at 37°C for 1 hour. The recombined lentiviral vector was transformed into One Shot® Stbl3™ cells as described in section 2.4.7 and the plasmid DNA was isolated as described in Section 2.4.9. The final lentiviral vectors were validated using sequencing (section 2.4.10).

2.8.2 LENTIVIRAL OVER EXPRESSION VECTOR CONSTRUCTION

The pLVX tight PURO lentiviral (Clontech, UK) vector was used for gene over expression in mammalian cells. The gene sequences were generated from image clones using specifically designed primers (Table 2.4), carried out by PCR (section 2.4.2). The PCR reaction products were ligated into the PCR2.1 vector using the TOPO TA Ligation Kit (Life Technologies, UK) as per the manufacturer's instructions. The vectors were propagated by transformation into DH5α cells (section

2.4.7) and the DNA was isolated using the QIAGEN mini prep kit (section 2.4.9). The gene sequences were verified using sequencing (section 2.4.10).

The GOI sequences were digested from the PCR2.1 entry vector using specific combinations of restriction endonucleases as described in section 2.4.4 and ligated into the pLVX vector (section 2.4.5). The restriction enzymes used were as follows:

SUMO	Bam HI and NotI
2A	NotI and MluI
Ubc9	MluI and EcoRI
HNF4 α	EcoRI

The ligated vector was transformed into Stabl3 cells as instructed in section 2.4.7 and subsequently purified (section 2.4.9). The final over expression vectors were validated using sequencing (section 2.4.10).

2.8.3 LENTIVIRAL PACKAGING

293FT cells were co-transfected with the ViraPower™ Packaging mix (Life Technologies, UK), containing the plasmid vectors that express the genes required for efficient viral packaging, and the pLenti4 lentiviral vector (LV). A no DNA and Lipofectamine® 2000 negative control was included in the experiment and was used to evaluate the results. 6×10^6 293FT cells were used for each a sample LV. For each transfection sample, DNA-Lipofectamine® 2000 complexes were prepared as follows: 9 µg of the ViraPower™ Packaging Mix was mixed with 3 µg of pLenti-based plasmid DNA (12 µg total) in 1.5 ml of Opti-MEM® I Medium without serum and was gently mixed. In a separate sterile 5 ml tube, Lipofectamine® 2000 was diluted by adding 36 µl into 1.5 ml of Opti-MEM® I Medium without serum, mixed

gently and incubated for 5 minutes at room temperature. The DNA was added to the diluted Lipofectamine® 2000 solution, mixed gently and incubated for 20 minutes at room temperature to allow the DNA: Lipofectamine® 2000 complexes to form. 293FT cells were trypsonised, counted and resuspended at a density of 1.2×10^6 cells/ml in Opti-MEM® I Medium containing serum. The DNA-Lipofectamine® 2000 complexes were added to a 10 cm tissue culture plate containing 5 ml of Opti-MEM® I Medium containing serum. 5 ml of the 293FT cell suspension (6×10^6 total cells) was added to the plate containing media and DNA-Lipofectamine® 2000 complexes and mixed gently by rocking the plate back and forth. The cells were incubated overnight at 37°C in 5% (v/v) CO_2, 95% (v/v) air. After 24 hours, the media containing the DNA-Lipofectamine® 2000 complexes was removed and replaced with complete culture medium. The virus-containing supernatants were harvested 48–72 hours post transfection by removing medium into a 15 ml sterile, capped, conical tube. The supernatants were centrifuged at 3000 rpm for 5 minutes at 4°C to pellet cell debris and filtered using a 0.45 µm low protein-binding filter to further purify the viral supernatants. The viral supernatants were aliquoted in 1 ml volumes and stored at -80°C for later use.

2.8.4 LENTIVIRAL TITRATION

Viral titres are a measure of the number of infective viral particles present in the packaged supernatant. It allows the control of the number of integrated copies within the host genome and provides reproducible results. As such the viral titre was measured for each LV. The day before transduction (Day 1), 293FT cells were trypsinised and counted before they were plated into a 6-well plate generating 30–

50% confluency at the time of transduction. The cells were incubated 37°C in 5% (v/v) CO_2, 95% (v/v) air overnight. Twenty four hours later, the lentiviral stock was thawed and 10-fold serial dilutions ranging from 10-2 to 10-6 were prepared. For each dilution, the lentiviral construct was diluted into complete culture medium to a final volume of 1 ml. The culture medium was removed from the cells and the diluted virus were gently mixed by inversion and added to one well of cells. 6 µg/ml of Polybrene® was added to each well to improve transfection efficiency. The plate was swirled and incubated at 37°C in 5% (v/v) CO_2, 95% (v/v) air overnight. The media containing virus was removed 24 hours later and replaced with 2 ml of complete culture medium. Twenty four hours later, the cells were trypsinised and the entire amount was re-plated into one 10 cm plate containing complete culture medium with the appropriate antibiotic to select for stably transduced cells. The medium was replaced with fresh medium containing antibiotic every 2–3 days. After 10–12 days of selection, the medium was removed and the cells were washed twice with PBS. 10 ml of crystal violet solution was added and incubated for 10 minutes at room temperature. The crystal violet stain was removed and the cells were washed twice with PBS. The blue-stained colonies were counted and the titre of the lentiviral stock was measured.

2.8.5 Antibiotic Kill Curves

Kill curves provide an accurate measure of the optimal concentration of the antibody required for cell selection. The respective cells were plated at approximately 25% confluence, 6 wells for each antibiotic was set up. The cells were left over night at 37°C in 5% (v/v) CO_2, 95% (v/v) air to allow them to adhere. Twenty four hours

later the media was changed and varying concentrations of the antibiotic were added to the media. The media was changed every other day and the percentage of surviving cells was observed. This was carried out for 14 days and the optimal concentration of the antibiotic was determined.

2.8.6 MAMMALIAN CELL TRANSFECTION

The cells of choice were plated in complete growth media as appropriate, for optimal transfection the cells were maintained at 30% confluency when possible. The Lenti4/BLOCK-iT™-DEST lentiviral stock was thawed and added into fresh complete medium at an MOI between 1 and 5, and was diluted if required. The total volume of medium containing virus was kept as low as possible to maximize transduction efficiency. Polybrene® was added to a final concentration of 6 µg/ml and the plate was swirled gently to mix. The cells were incubated at 37°C in 5% (v/v) CO_2, 95% (v/v) air overnight. The medium containing the virus was removed and replaced with fresh media 24 hours later. The cells were incubated at 37°C in 5% (v/v) CO_2, 95% (v/v) air overnight. The cells were maintained in the appropriate media and the RNA and protein extracts were collected at specific time points.

CHAPTER THREE

DEFINING HUMAN EMBRYONIC STEM CELL CULTURE CONDITIONS AND THEIR SUBSEQUENT DIFFERENTIATION INTO HEPATIC ENDODERM

CHAPTER 3

DEFINING HUMAN EMBRYONIC STEM CELL CULTURE CONDITIONS AND THEIR SUBSEQUENT DIFFERENTIATION INTO HEPATIC ENDODERM

3.1 INTRODUCTION

3.1.1 PROPERTIES AND CHARACTERISTICS OF hESCs

Human embryonic stem cells (hESCs) have the potential to provide an inexhaustible supply of a variety of human somatic cell types (Anna et al., 2005). These in turn can be used for applications such as disease modelling, drug discovery and cell therapy. hESCs can also be utilised for deciphering complex mechanisms throughout development and during organ formation (Vazin et al., 2010). Two properties make them ideal candidates for modelling and therapeutic applications; this includes their ability to indefinitely self renew and to differentiate into cell types from all three germ layers; the mesoderm, ectoderm and endoderm (Pera et al., 2000, Henrik, 2005).

hESCs are isolated from the inner cell mass of blastocyst stage embryos {Thomson, 1998, Peura, 2007). Once an egg has been fertilized and a diploid zygote has formed during early embryogenesis, the blastocyst is generated. The blastocyst has two layers; the outer layer referred to as the trophoblast and the inner layer defined as the

embryoblast, also known as the inner cell mass (ICM), which gives rise to the embryo. (Peura et al., 2011) The ICM is then dissociated from the embryo using either immunosurgery, mechanical or enzymatic dissection and is subsequently cultured in vitro using a variety of methods. The latter form of dissection is favoured as it prevents the contamination of animal derived products, maintaining the derived cells at a suitable grade for clinical applications.

hESCs are characterised using an array of stem cell specific markers including; the expression of ES cell specific transcription factors; Octamer 4 (Oct 4) and Nanog, their cell surface expression of stage specific embryonic antigens (SSEA) 3 and 4, Tra 160 and 181 (Xu et al., 2001, Amit et al., 2006, Anna et al., 2005) and their morphology (Henrik, 2005). hESCs typically form tightly packed colonies with defined edges. The population of cells have a high nucleus to cytoplasm ratio with pronounced nucleoli (Pera et al., 2000). Karyotyping is frequently performed to ensure the correct chromosome number is present and that no major transpositions occurred (Campos et al., 2009). DNA microarrays, short tandem repeat analysis, fluorescent in situ hybridization and whole genome single nucleotide polymorphism can also be used to confirm the precise hESC signature within the population (Josephson et al., 2006, Brimble et al., 2004, Mitalipova et al., 2005). hESC pluripotency is measured by their ability to form embryoid bodies and subsequently spontaneously differentiate into cell types from all three germ layers (Hannoun et al., 2010, Barbara S Mallon et al., 2006, Cai et al., 2006). Each germ layer is identified using a specific marker; the endoderm is usually defined by the presence of alpha feto protein (AFP), the ectoderm stains positive for β-tubulin and the mesoderm

usually contains smooth muscle actin (SMA) (Hannoun et al., 2010). hESCs can also be cultured in vitro for over 100 passages and have been directly differentiated into a variety of cell types such as; cardiomyocytes, neurons, epithelial cells and hepatocytes (Hannoun et al., 2010, Niebruegge et al., 2009, Van Hoof et al., 2009, Trounson, 2006, Hay et al., 2008).

3.1.2 hESC Culture Conditions

hESCs are frequently maintained on mouse embryonic fibroblasts (MEFs) in the presence of serum replacement media and basic fibroblast growth factor (bFGF) (Barbara S Mallon et al., 2006, Henrik, 2005, Jean et al., 2006). The MEF feeder cells are able to supply the essential factors and cell attachment support that are required for healthy maintenance of undifferentiated pluripotent hESCs (Thomson et al., 1998). In an attempt to move away from the use of xeno related products in hESC culture, MEFs have been replaced by a variety of human derived fibroblasts and fibroblast like cells; ranging from fetal foreskin cells (Amit et al., 2003) to bone marrow cells (Cheng et al., 2003); all of which supported viable hESC growth.

Due to the continuous requisite of using hESCs in a pure population and in a scalable manner, feeder free conditions were developed. hESCs were cultured on the basement membrane, Matrigel™; a membrane preparation extracted from a murine Englebreth-Holm-Swarm sarcoma (Xu et al., 2001, Amit et al., 2004), which provided the necessary attachment support for cell growth. Other extracellular matrices that have been successfully employed include fibronectin, vibronectin and laminin or a combination of the three (Braam et al., 2008, Amit et al., 2006, Rodin et

al., 2010). Initially, MEF conditioned media (CM) was used to provide the required growth factors for healthy hESC maintenance on various extracellular matrices (He et al., 2005, Xu et al., 2001). Despite its advantages, CM culture systems suffer from lack of definition, batch to batch variability, presence of animal components and the labour intensity associated with CM production; all of which result in a poorly defined cellular model (Hannoun et al., 2010).

3.1.3 hESC Culture Definition

In an attempt to create more defined conditions, a number of laboratories executed the characterisation of the essential components of maintenance media and extra cellular support matrices. Comparative proteomic analysis was performed on media before and after conditioning by various fibroblasts to try and isolate the important factors required for hESC maintenance. A total of 34 proteins were regarded as significant as these were found in media conditioned by MEFs, human fetal and neonatal fibroblasts (Prowse et al., 2007). Overall various studies have confirmed the following proteins to be essential for hESC culture; bFGF (Wang et al., 2005), Activin A (Beattie et al., 2005), transforming growth factor beta-1 (TGF β-1) (James et al., 2005) and Noggin (Pera et al., 2004); used to repress the bone morphogenic protein (BMP). These factors were added to serum free media in various concentrations and combinations, depending on the research group. Overall, the correct signalling pathways required for the adequate maintenance of hESCs in a pluripotent non differentiated state seemed to be activated within these systems.

Chin and colleagues successfully managed to replace MEFs/CM with a cocktail of recombinant proteins including; plasminogen activator inhibitor, monocyte chemoattractant protein 1, insulin like growth factors binding proteins 2 and 7, interleukin 6 and pigment epithelium derived factor; supporting healthy growth of hESCs cultured on fibronectin (Chin et al., 2007). Another study has developed a defined media (RegES) containing only recombinant, synthetic or human derived components capable of maintaining hESCs in an undifferentiated state. The main ingredients consist of knock out Dulbecco's modified Eagle's medium (serum free basal media), human serum albumin, vitamins, antioxidants, trace minerals, amino acids and growth factors (Rajala et al., 2010). hESCs cultured in RegES maintained the expression of the ES markers Oct 4, Nanog and Sox 2 in conjunction with the expression of cell surface antigens SSEA 4 and Tra-160. RegES cultured hESCs retained the correct colony morphology and full pluripotency, indicated by forming cell types from all three germ layers via EB formation. Directed differentiation into cardiomyocytes and neural cell types was also achieved. (Rajala et al., 2010)

To help standardise the culturing techniques of hESCs, various companies have released defined xeno free media into the market. The most popular media used to date are StemPro (Invitrogen), mTeSR 1 (Stem Cell Technologies), HEScGro (Millipore), NutriStem (Stemgent) and X-Vivo (Lonza). Research groups internationally have carried out trials to select the most effective media for culturing hESC using a number of different hESC lines and extra cellular matrices to help create the ideal environment (Chin et al., 2010, Melkoumian et al., 2010, Sugii et al., 2010, Hannoun et al., 2010). Despite claims that all of the above media are capable

of maintaining hESCs in vitro; the majority of the publications seem to have shown mTeSR and StemPro to be the most effective, including our own research, discussed in section 3.2. Recently, the new fully defined E8 media has been shown to promote efficient self-renewal and differentiation of hESCs and iPSCs (Chen et al, 2011).

3.1.4 METHODS OF HESC EXPANSION

Another aspect of hESC culture is their expansion. The rate of hESC growth is highly dependant on the culture conditions and the hESC line itself. Usually at 80-95% confluency, the cells are disaggregated either mechanically or enzymatically; using collagenase, dispase and trypsin, all of which are animal derived making the cells unsuitable for the clinic (Hoffman et al., 2005). Recombinant human collagenase has been successfully utilised and has resulted in the improvement of the culture conditions towards good manufacturing practice (GMP) standards (Crook et al., 2007). Data has shown both methods to be successful in individual cases; however there is variation between the results. Mechanical passaging is labour intensive and can not be accurately reproduced. On the other hand, enzymatic dissection has also shown to result in genetic abnormalities within the cells (Brimble et al., 2004, Mitalipova et al., 2005). An automated approach is required to provide a reproducible and efficient method of hESC expansion whilst maintaining hESC pluripotency and karyotypic stability (Mitalipova et al., 2005).

3.1.5 IMPORTANCE OF THE HESC MICROENVIRONMENT

It is vital to recapitulate the surroundings of hESCs in vitro as observed in the in vivo niche. As such, cell: cell and cell: matrix interactions play an important role in hESC

culture and differentiation, mediated by specific integrin expression. Most hESC culture methods rely on two dimensional (2D) systems; however studies have shown that culturing hESCs in a three dimensional environment (3D) enhances cell proliferation and self renewal as it more accurately mimics the in vivo environments (Lee et al., 2007, Postovit et al., 2006). The hypothesised mechanism suggests the requirement for physical and mechanical cues in hESC growth. This has been shown by the activation of the GTPase Rac and the subsequent induction of the phosphoinositide 3-kinase (PI3K) pathway, resulting in the up regulation of Nanog expression, required for hESC self renewal (Nur-E-Kamal et al., 2006).

A number of cases have successfully derived defined 2D matrices for the maintenance of hESCs. Melkoumian and colleagues synthesised a peptide- acrylate surface (PAS) whereby biologically active peptides; vibronectin, laminin, fibronectin and bone sialoprotein, required for hESC self renewal are conjugated to an acrylate surface (Melkoumian et al., 2010). The hESCs cultured on PAS in the presence on X Vivo 10 media supplemented with bFGF and TFG-1 retained all ES cell specific properties, as stated in section 3.1.1 (Melkoumian et al., 2010). The cells could subsequently be differentiated into functional cardiomyocytes in a scalable manner (Melkoumian et al., 2010). Similar observations were found when using another synthetic polymer, poly [2-(methacryloyloxy) ethyl dimethyl-(3-sulfopropyl) ammonium hydroxide] (PMEDSAH) (Villa-Diaz et al., 2010). PMEDSAH supported long term self renewal of undifferentiated hESCs with retained pluripotency when cultured in a number of media; CM, human cell conditioned media (hCCM), mTeSR and StemPro (Villa-Diaz et al., 2010). Although successful 2D systems are

extensively used, it is probable that scalable and defined 3D culture systems will provide a method for large scale and cost effective manufacturing in the future. Such a scheme is required for the widespread use of hESCs as tools in research and cell based therapies in the clinic.

3.1.6 EXPRESSION PROFILING OF HESCS

Recent research has established the relationship between stem cell lines derived from early mouse and human embryos. The gene expression profiles of hESCs more closely resemble that of the epiblast over the inner cell mass, suggesting the epiblast nature of hESCs (Reijo Pera et al., 2009). The epiblast is formed at a later stage in development and is derived from the inner cell mass. Hence, hESCs with epiblast like properties are usually defined as cells generated from a later stage in development. This is further supported by data showing similarity between mouse epiblast stem cells and hESCs, both on gene expression and phenotypic levels (Reijo Pera et al., 2009). On the other hand, mouse embryonic stem cells are fixed at an earlier stage in development and are more comparable to the inner cell mass (Brons et al., 2007, Tesar et al., 2007). However, recent studies have shown the importance of the microenvironment and culture conditions on defining the epiblast or pre-epiblast stage of hESCs and hence their expression profiles (Tavakoli et al., 2009, Allegrucci et al., 2007). It has also been noted that there are large differences in the gene expression profiles of different hES cell lines, further supporting the above hypothesis (Reijo Pera et al., 2009, Allegrucci et al., 2007). The majority of differences associated between the various hESC lines can be attributed to epigenetic changes due to the surrounding environment (Brink et al., 2008). Extensive

characterisation of hESC lines has resulted in the isolation of genes affected by genetic imprinting via epigenetic regulation (Brink et al., 2008). Various epigenetic changes including distinctive methylation patterns and divergent histone modifications have been detected in hESC lines (Frost et al., 2011). Studies have accredited these differences to inconsistencies in the culture environments (Frost et al., 2011).

During the past year, three groups have managed to alter the state of hESCs to behave more like mouse embryonic stem cells by activating reprogramming transgenes such as KLF4 (Hanna et al., 2010), or by the addition of LIF, a factor essential for mESC maintenance and other small molecules (Xu et al., 2010) or a combination of both (Buecker et al., 2010). The above data substantiates the dynamic and fluid nature of hESCs, an important property that should be considered when defining ideal culture systems.

The above section summarises the important points associated with hESCs and the various culture techniques used to maintain the cells in an undifferentiated pluripotent state. The main challenges in defining hESC culture are integration of the media with the matrix support, scalability, reproducibility and limited knowledge of the underlying mechanisms. The next section of the chapter highlights the investigation carried out to help overcome the above issues. The employment of a defined, serum free commercially available media will improve the reproducibility of the culture system as it reduces batch-to-batch variation and allows for scalability.

Furthermore, this system could be implemented as an accurate model for understanding the mechanisms employed by hESCs.

3.2 RESULTS

DEFINING HUMAN EMBRYONIC STEM CELL CULTURE CONDITIONS

A large amount of time and effort has been invested in creating the ideal culture conditions for maintaining hESCs in a defined system. However, despite a number of successful approaches; a standardised, reproducible, scalable and efficient method has yet to be determined. This is the first step to providing a platform required to realise the full potential of hESCs. As such, a media trial was carried out to further refine hESC culture techniques. A more defined serum-free commercially available media, mTeSR (MT) (Ludwig et al., 2006) was employed to further standardise hESC maintenance using CM as the base control.

3.2.1 CULTURE AND CHARACTERISATION OF HESCS MAINTAINED IN CM AND MT

hESCs previously cultured in CM were gradually transitioned into MT over 14 days. At the first passage hESCs cultured in 100% CM, were transferred into a combination of CM and MT with an 80:20 ratio. At the second passage, the cells were cultured in a 60:40 ratio of CM and MT, respectively. At each following passage the cells were cultured in a mixture of CM:MT at the following ratios; 40:60, 20:80 and finally 100% MT. The hESCs were then maintained in both CM and MT for a further 30 passages and then characterised for their ES cell identity and pluripotency accordingly. hESC identity was assessed by their morphology, the identification of transcription factors ascribed to pluripotency; Oct4 and Nanog, and

their cell surface expression of stage specific embryonic antigens (SSEA) 3 and 4 and Tra 16-0 and 1-81 (Hay et al., 2008). Furthermore, their pluripotency was measured by their ability to spontaneously differentiate into all three germ layers; the mesoderm, ectoderm and endoderm, and directly differentiate into hepatocytes using a standardised protocol (Hay et al., 2008, Fletcher et al., 2008), an elaborate description can be found in section 3.2.3.

The first stage of assessing hESC identity was to compare the morphology between hESCs grown in CM and MT. hESCs grown in CM formed compact colonies with slightly loose edges, and little spontaneous differentiation was observed (Figure 3.1 A, arrows). The colony morphology was maintained over 30 passages (Figure 3.1 a-f, arrows). However, hESCs maintained in MT adopted a morphology similar to hESCs cultured in CM for approximately 10 passages (Figure 3.1 B, g and h). The hESCs then formed highly defined tightly packed dome-like colonies with no observable spontaneous cellular differentiation (Fig. 3.1 B, i-l, arrows). hESCs cultured in both media had a high nuclear to cytoplasm ratio with very defined nuclei.

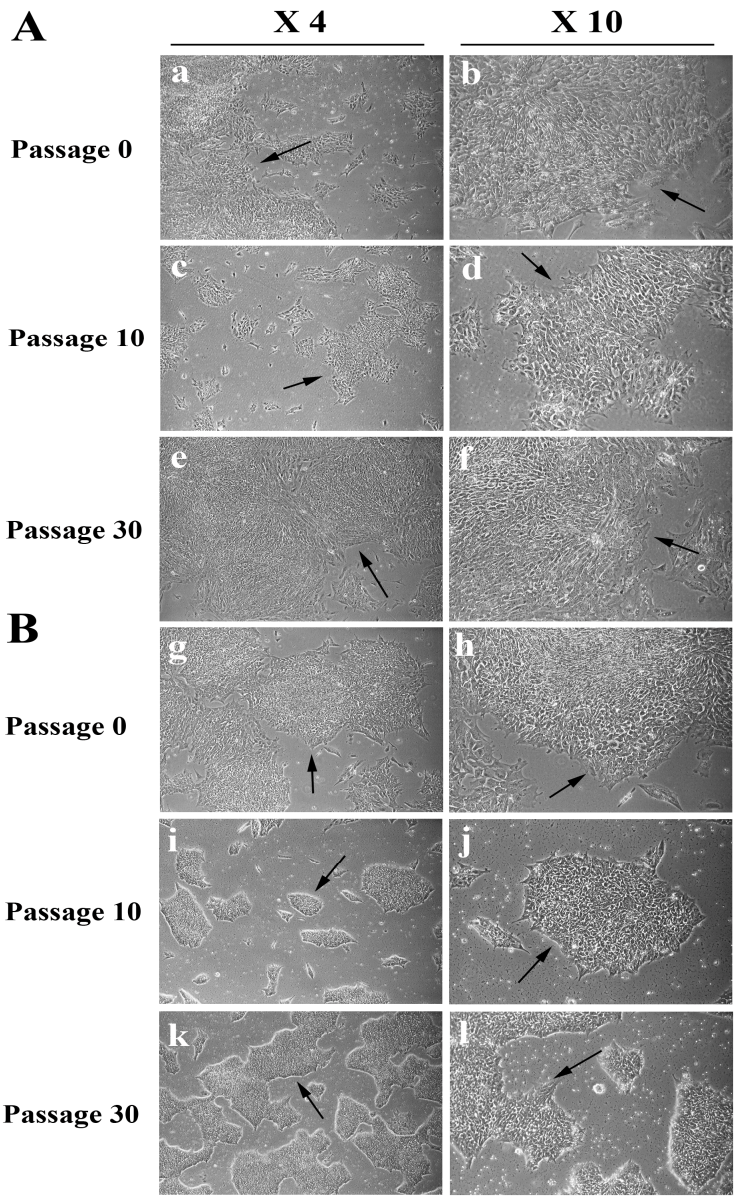


Figure 3.1 – Morphological Analysis of hESCs cultured in CM and MT.

hESCs cultured in various media have been known to adopt various morphological properties. **A** displays hESCs cultured in conditioned media from 0 (a and b), 10 (c and d) and 30 (e and f) passages. As observed, there are no significant changes in the morphology of the colonies for over 30 passages, **A**. hESCs cultured in CM form defined colonies with slightly loose edges as seen in b, d and f (arrows). Little differentiation is found between the colonies. The hESCs maintain a large nucleus to cytoplasm ratio with defined nuclei. **B** depicts hESCs cultured in mTeSR. Morphological changes are noted between passages 0 and 10. At passage 0 (g and h) the cells adopt a CM like morphology with defined colonies and loose edges. However, at passage 10 (i and j) the colonies pack in to form more dome like structures with fully defined edges as seen in h, j and l (arrows). The hESCs form more circular shaped colonies. This morphology is maintained for 30 passages (k and l). As with hESCs cultured in CM, MT cultured hESCs have high nuclear to cytoplasm ratios with defined nuclei and little to no differentiation observed between the colonies.

Chromosomal abnormalities have been attributed to long-term culture of hESCs and the likelihood may be related to the media conditions employed. As a result, we carried out karyotypic analysis on hESCs cultured in CM and MT to ensure chromosome integrity. Forty-eight hours after the hESCs were passaged; Colcemid (GIBCO, UK), a drug used to arrest cells at metaphase, was added to the cells, which were then harvested 1-2 hours later. Chromosome slides were then prepared and Giemsa stained (Section 2.3.3), a technique used to band the chromosomes. Figure 3.2 confirms the correct number of paired chromosomes in hESCs cultured in both CM (a) and MT (b) with no detectable abnormalities (n=10). In conclusion, the hESCs cultured in CM and MT maintained a normal karyotype for over 30 passages, indicating the stable nature of the environment.

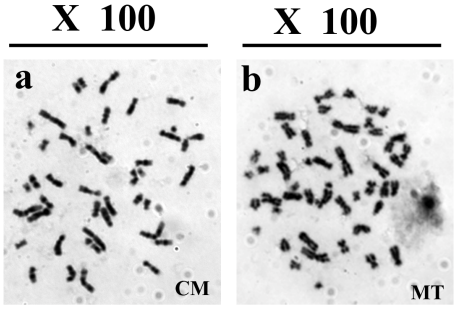


Figure 3.2 – Karyotypic Analysis of hESCs cultured in CM and MT.
Chromosomal integrity is an important property required for healthy maintenance of hESCs especially during long-term culture. As such, after 30 passages CM and MT cultured hESCs were karyotyped to analyse their chromosomal structures, this was carried out using Giemsa banding. The pictures display the chromosomes from hESCs cultured in CM (a) and MT (b). It can be noted that hESCs cultured in both MT (b) and CM (a) have the correct number of paired chromosomes with no detectable abnormalities (n=10). The images were captured at x100 magnification on a Leica Light microscope. Chromosome analysis was carried out in house.

Further to morphological and karyotypic analysis, we investigated the stem cell nature of the hESCs cultured in both media. hESCs cultured in both CM and MT stained positive for Oct 4 protein expression as depicted by the immuno-fluorescence (Figure 3.3 A, a and b) and Nanog gene expression detected by PCR (Figure 3.3 B). hESCs cultured in MT displayed higher levels of Nanog expression (Figure 3.3 B)

which may correlate to a more homogenous stem cell population than that observed in hESCs cultured in CM. The high level of expression of Oct4 and Nanog, two main transcription factors associated with hESC 'stemness', confirmed the undifferentiated state of the hESCs whilst maintaining pluripotency; which will be analysed in more detail in section 3.2.2 and 3.2.3.

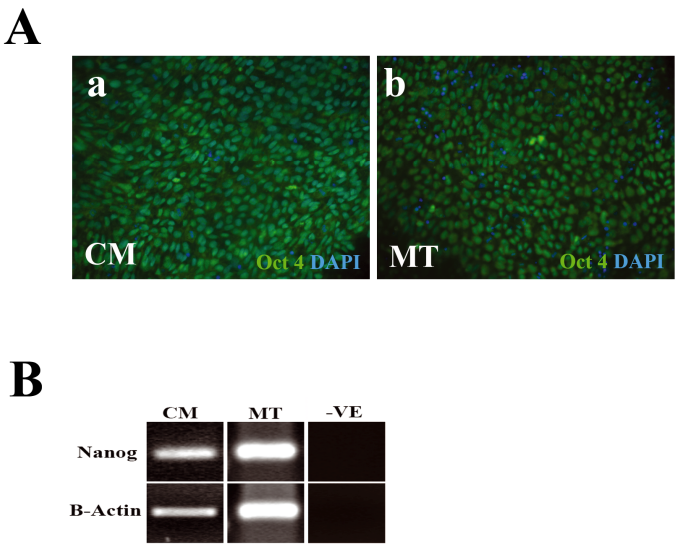


Figure 3.3 – Characterisation of hESCs cultured in CM and MT.
A hESCs cultured in both media are Octamer 4 (Oct4) positive as shown in conditioned medium (CM, panel a) and mTeSR (MT, panel b) at x20 magnification using a Nikon TE3000/U inverted microscope. This supports the efficient maintenance of hESCs by both MT and CM, as the hESCs expressed stem cell specific marker Oct4. **B** indicates the expression of the hESC transcript Nanog (band at correct size), further supporting the stem cell status of hESCs maintained in both different media. B actin is used as a loading control and a negative RT was included

Fluorescence activated cell sorting (FACS) was used to examine the expression of cell surface antigens, another technique employed to characterise the stem cell identity of hESCs. The surface markers used to identify hESCs within a population include SSEA 3 and 4 in conjunction with Tra 1-60 and 1-81. The Tra surface antigens have been suggested to be the more sensitive markers, as their expression decreases almost instantaneously when hESCs begin to differentiate. The results

displayed in figure 3.4 show background level of SSEA-1 expression (a differentiation marker) in both CM and MT cultured hESCs, as expected.

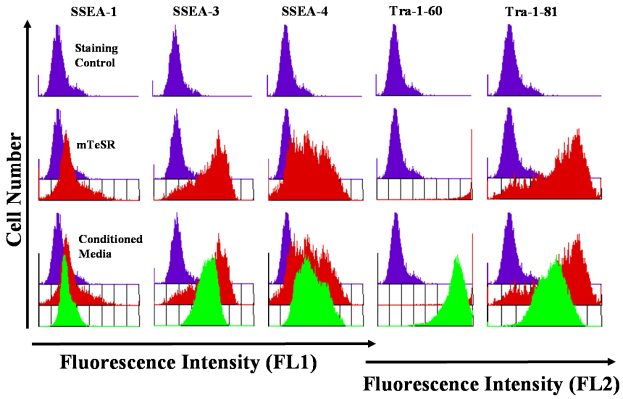


Figure 3.4 – Investigation of Surface Marker Expression of hESCs cultured in CM and MT.
The FACS plots show hESC surface marker expression levels as expected, including stage specific embryonic antigens (SSEA) and Tumour rejection antigens (Tra). SSEA 3, 4, Tra 1-60 and 1-81 are expressed at 92.11%, 27.26%, 90% and 98.62% in hESC cultured in CM and; 97.72%, 46.38%, 89.60% and 79.59% in hESCs cultured in MT; respectively. Negligible levels of SSEA-1 were detected further supporting the stem cell nature of these hESCs; 4.8% in CM and 18.2% in MT. A negative staining control, secondary antibody only, was used to define levels of background staining.

The levels of SSEA 3, Tra 1-60 and 1-81 expression was confirmed in greater than 80% of the cells, confirming the undifferentiated state of the hESCs. It should be noted that the levels of SSEA-4 were lower than expected, but was to be due to the inefficiency of the antibody and not the hESCs themselves. In addition to the increased Nanog expression in hESCs cultured in MT, a higher level of Tra 1-60 expression was observed, both correlating to the more stem cell nature of the cells. SSEA 3, 4 and Tra 1-81 expression are comparable between both media. SSEA 3, 4, Tra 1-60 and 1-81 were expressed in 92.11%, 27.26%, 90% and 98.62% of hESCs cultured in CM and; 97.72%, 46.38%, 89.60% and 79.59% of hESCs cultured in MT; respectively.

In conclusion, hESCs cultured in both CM and MT exhibited the required characteristics that define the undifferentiated and pluripotent state of the cells. Both media could sustain hESCs self-renewal for over 30 passages maintaining the correct morphology (Figure 3.1), chromosome stability (Figure 3.2) and Oct 4 and Nanog expression (Figure 3.3). This data was additionally supported by the expression pattern of the cell surface antigens, refining their undifferentiated state. MT successfully maintained the undifferentiated hESC self renewal for over 30 passages, comparable to CM, and is therefore able to provide a stable environment for hESC culture. The next two sections will focus on the pluripotent properties of the hESCs cultured in CM and MT.

3.2.2 Investigating Pluripotency of hESCs Culture d in CM and MT

Embryoid body (EB) formation is a standardised method for measuring hESC pluripotency (Xu et al., 2001). The technique assesses the ability of hESCs to differentiate into cell types from all three germ layers; the mesoderm, endoderm and ectoderm, which occurs spontaneously. The cell types associated with the germ layers are frequently identified using the following antibodies: alpha-feto protein (AFP) defines endodermal cells, smooth muscle actin (α SMA) depicts mesoderm and β tubulin III categorizes ectoderm. hESCs at 80-90% confluency were lifted and placed into a suspension suitable for promoting cell aggregation (Section 2.2.2) (Figure 3.5 A). After 7 days the EBs generated were well defined and vacuolated (Figure 3.5 B), suggesting efficient EB formation. The EBs were then re-plated on gelatin-coated wells and were allowed to spontaneously differentiate over a 14-day period. Two weeks provides adequate time to allow the hESCs to generate cell types

from all three germ layers. The resulting cell types are then fixed in ice cold methanol and are stained with antibodies that are specific to each individual lineage (Section 2.3.1) (Figure 3.6).

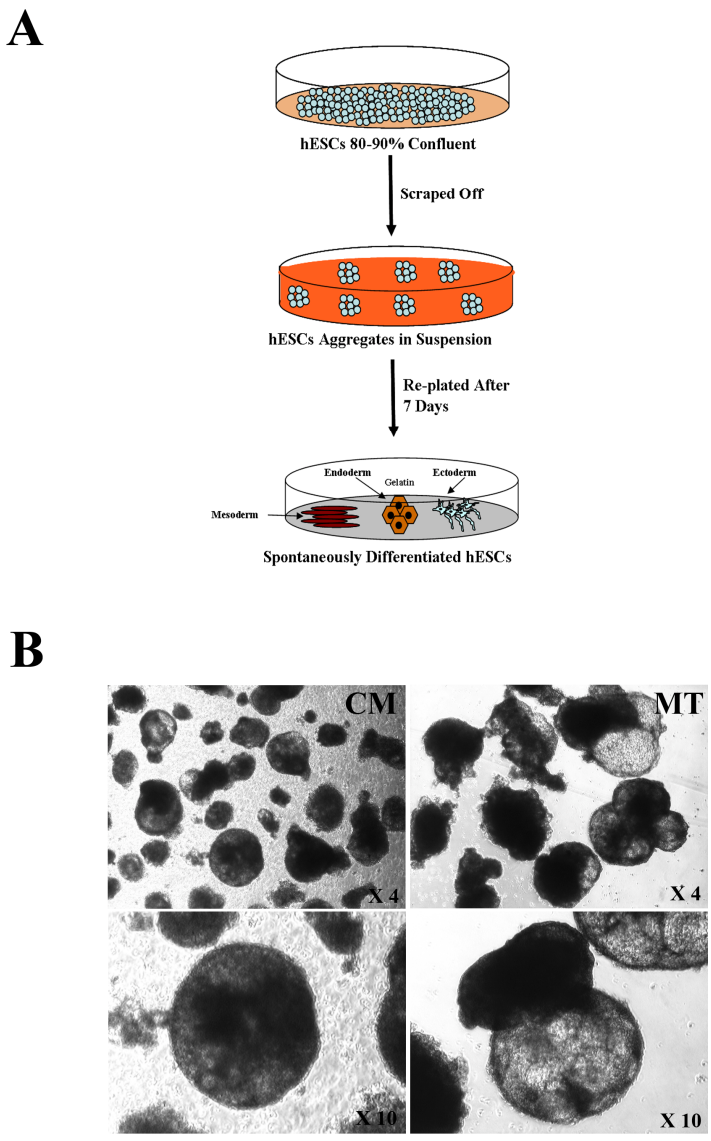


Figure 3.5 – Investigation of pluripotency in hESCs cultured in CM and MT.
A. Embryoid body formation is one method of investigating hESC pluripotency. hESCs at 80-90% confluency cultured on Matrigel™ are scraped off in large clumps and are allowed to aggregate in low culture attachment plates for 7 days in EB specific media, see material and methods for details. The EBs are then plated onto gelatine coated plates for a further 14 days where they spontaneously differentiated into cell types of all three germ layers; mesoderm, endoderm and ectoderm. The formation of these cell types indicated sufficient hESC pluripotency. **B.** These images display embryoid body formation by hESCs cultured in CM and MT at x4 and x10 magnifications. The EB's are well defined and vacuolated, suggesting efficient EB formation in both media. The images were captured using a Nikon TE3000/U inverted microscope.

Figure 3.6 displays the resulting cell types formed following EB spontaneous differentiation from hESCs cultured in CM and MT. The hESCs formed cell types that were positive for all three germ layers; AFP (endoderm), α SMA (mesoderm) and β tubulin III (ectoderm), and are comparable between both media. The cellular

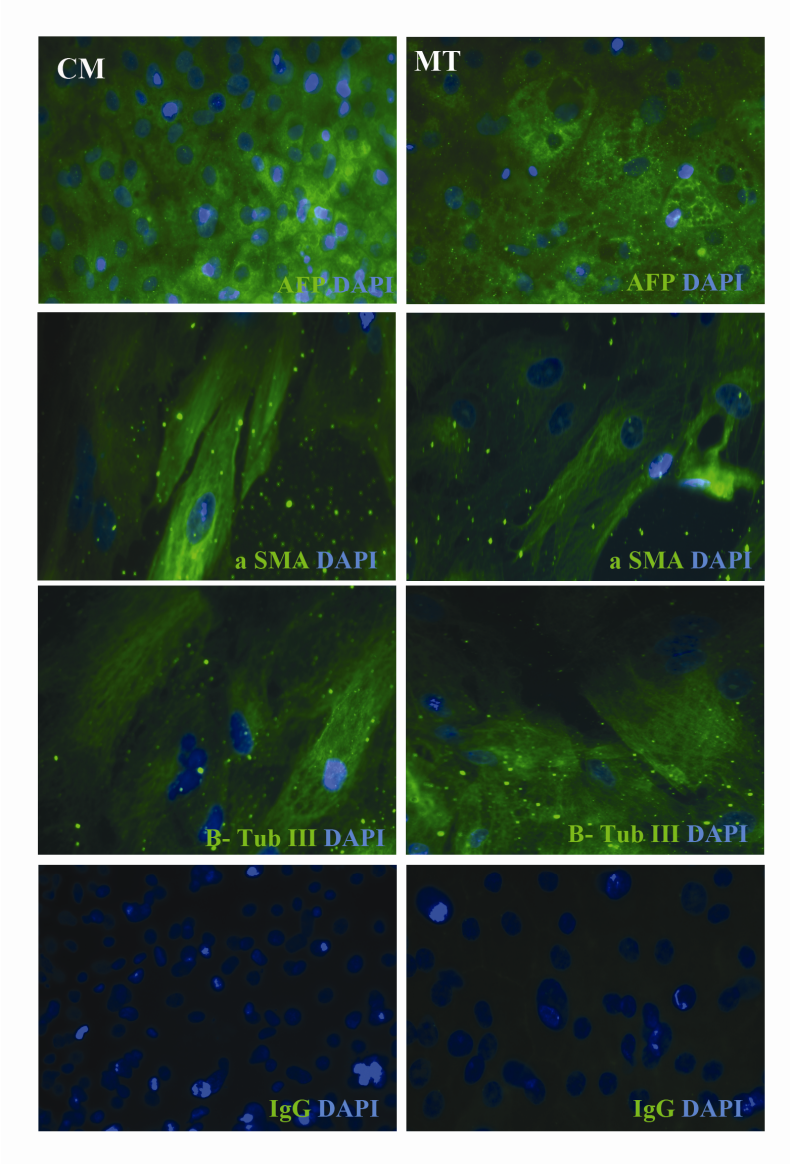


Figure 3.6 – Analysing the pluripotency of hESCs cultured in CM and MT.
Immunocytochemical analysis shows that the hESCs are able to form cell types from all three germ layers after being cultured in CM and MT, indicative of their pluripotent abilities. AFP positive cells denote cells from the endoderm lineage, muscle actin is used to define mesodermal cells and beta tubulin III defines cells from the ectodermal lineage. These images were captured at x40 magnification using a Leica DMIRB inverted fluorescent microscope, all images are merged DAPI and FITC images, with IgG used as a negative control. Abbreviations: AFP – alpha-fetoprotein, α-SMA – alpha smooth muscle actin and β Tub III – Beta Tubulin III.

morphologies also corresponded to the germ layer formed; for example, cells positive for α SMA displayed streaky sheet like structures that are normally associated with muscle. This data demonstrates that the hESCs cultured in MT remain pluripotent for over 30 passages and are able to spontaneously differentiate into all three germ layers in a manner comparable to hESCs maintained in CM.

3.2.3 Directed Differentiation and Characterisation of hESC Derived Hepatic Endoderm

Efficiently directing hESC differentiation to the appropriate cell type is essential if stem cell-derived products are to be manufactured cost effectively. hESCs maintained in CM and MT were scaled up to the desired quantity and hepatic differentiation was driven using Activin A andWnt3a stimulation, Section 2.2.3. This protocol was employed over the Activin A and Sodium Butyrate method as it was more efficient at generating HE and displayed improved function specifically in ureagenensis and serum protein production; including proteins AFP and fibrinogen (Hay et al., 2008).

The differentiation protocol mimics the stages observed throughout development, whereby the hESCs are primed towards definitive endoderm, followed by hepatic specification and maturation. Initially the hESCs are cultured in RPMI/B27 supplemented with Activin A and Wnt3a, this stage primes the cells towards definitive endoderm. The cells are then cultured in SR/DMSO media, which induces hepatic specification. Subsequently, the cells are transferred to L15 supplemented with hepatocyte growth factor (HGF) and Oncostatin M (OSM) driving hepatic

maturation. The resulting hepatic endoderm (HE) is then characterised using morphological, transcriptional and functional analysis. This form of characterisation encompasses both the transcriptional and translational levels of the resulting HE, thus accurately defining the end product.

hESCs cultured in MT and CM were differentiated using the above protocol; RNA and protein samples were collected at various time points which were then used for further analysis. The first stage was defining the transcriptional gene patterning of the differentiating cells.

3.2.3.1 TRANSCRIPTIONAL ANALYSIS OF HESC DERIVED HE

We investigated the expression of a number of genes required in hepatic differentiation using SOLEXA (Cuddapah et al., 2009), a method of quantifying absolute levels of gene transcription, Section 2.4.11. The key time points examined during the analysis were day 0; undifferentiated pluripotent hESCs, day 3; definitive endoderm, day 10; hepatoblast like cells, day 17; differentiated hepatocyte like cells which were compared to fetal and adult human hepatocytes.

Figure 3.7 displays the expression pattern of stage specific genes involved in hepatocyte formation. As expected, peak Oct4 and Nanog (ESC specific transcription factors) expression levels were detected at day 0 and decreased as hESCs differentiated into primitive endoderm. No Oct4 and Nanog expression was detected in our control adult and fetal hepatocytes. As cells differentiated through the primitive streak towards definitive endoderm we detected the expression of

Brachyury followed by goosecoid and SOX 17 expression once definitive endoderm was generated. GATA 4 expression was noted at day 10 as the cells developed into foregut endodermal like cells, specifying hepatic commitment.

In conclusion, the gene patterning observed in our HE differentiation model is comparable to the expression pattern noted in the developing liver during embryogenesis. This supports that the differentiation model and the resulting HE are accurate representations of the in vivo system.

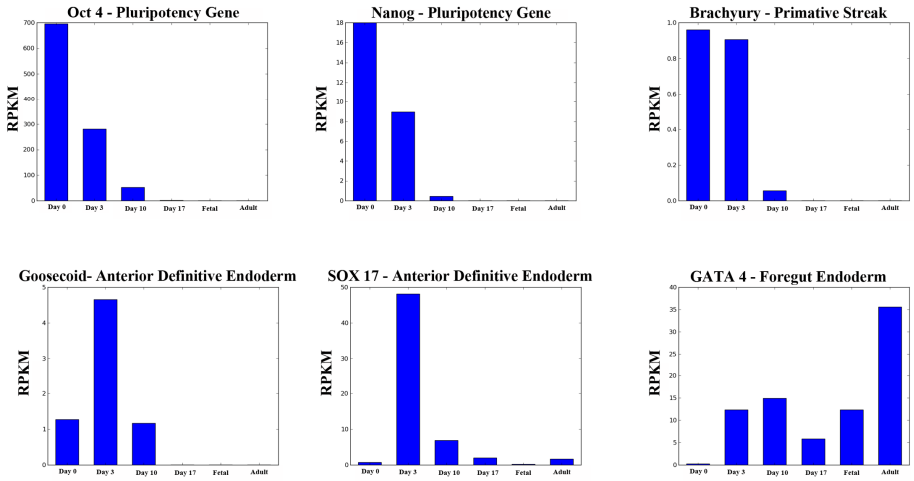


Figure 3.7 – Transcriptional Analysis of Hepatic Differentiation
In conjunction with analysing pathways at a translational level, SOLEXA was carried out to observe changes in gene expression at specific time points. Panel A depicts gene expression changes from day 0, 3, 10 and 17 hESC derived HE as well as fetal and adult human hepatocytes. At each stage of differentiation key genes were chosen and analysed. Oct 4 and Nanog were used to define the undifferentiated state of the cells (Nichols et al., 1998, Hoffman et al, 2005). Brachyury defined the primitive streak (Tam et al., 2003, Vaillancourt et al., 2009) followed by Goosecoid and SOX 17 that define the anterior definitive endoderm (Hannoun et al., 2010). As the cells commit to the hepatic lineage GATA 4 was used to define the foregut endoderm (Duncan, 2003, Bossard et al., 1998).

3.2.3.2 CHARACTERISATION OF HESC DERIVED HE

In addition to gene expression profiling, we noted the various morphological changes associated with hepatic differentiation in hESCs cultured in CM and MT, Figure 3.8 A (CM) and B (MT).

At approximately 20-30% confluency, hESCs cultured in CM and MT were directly differentiated to HE. Images were captured at days 0, 3 and 17 to note the morphological changes throughout hepatic differentiation. At day 0 both CM and MT cultured hESCs maintain the expected colony structure as described in section 3.2.1. As the cells are primed using Activin A and Wnt3a, they begin to proliferate and develop a more definitive endoderm like structure, day 3 (Figure 3.8 A and B).

The cells commit to hepatic specification, days 5-10, and being maturing, day 13. At day 17, the HE displays morphological traits associated with mature hepatocyte. The HE consisted of large hexagonal shaped cells with canaliculi like structures surrounding them. The hepatic nuclei are large and well defined, another property correlated to hepatic maturity. In conclusion, the morphological changes observed support hepatic maturation, and are comparable between both media.

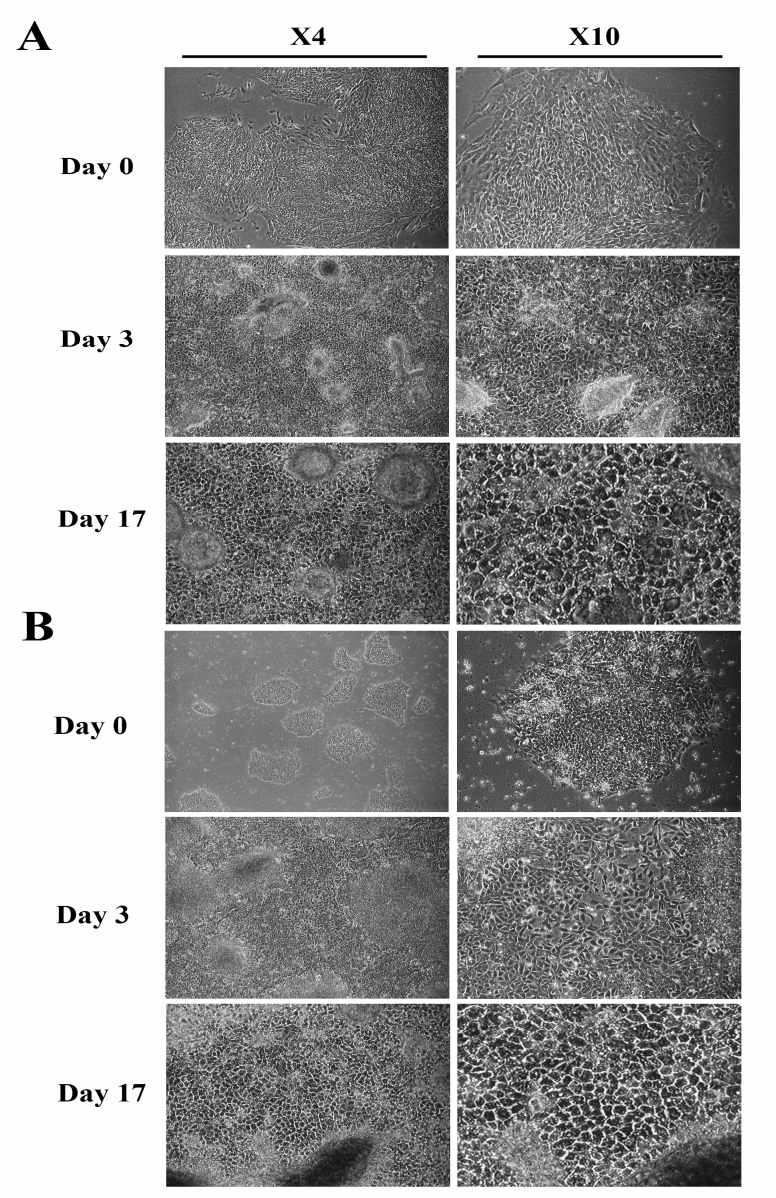


Figure 3.8 – Morphological Analysis of hESCs Derived HE, cultured in CM and MT.
Directly differentiating hESCs down the hepatic lineage is associated with significant morphological changes within the cells. Each morphological change is defined by a specific stage in the differentiation protocol. The images display the morphological alterations of hESCs cultured in CM (**A**) and MT (**B**) as they are differentiated into hepatic endoderm (HE). Similar patterns are observed between CM and MT derived HE. At day 0 the hESCs maintain normal ES cell morphology. Once stimulated to differentiate down the endodermal lineage, the cells adopt a definitive endoderm like shape at day 3. The differentiating cells then commit towards the hepatic lineage followed by maturation between days 15 and 17. The cells adopt a hexagonal shape and are surrounded by canaliculi like structures. These properties are similar to those observed in primary adult hepatocytes, suggesting successful differentiation into HE. The images were captured using a Nikon TE3000/U inverted microscope at x4 and x 10 magnification.

To further characterise our in vitro-derived HE, we assessed the repertoire of hepatic gene expression by RT-PCR (Section 2.4.2, 2.4.3). HE derived from hESCs cultured in MT expressed a number of hepatic transcripts; including HNF4 α, albumin and tryptophan oxidase (TO) indicative of hepatic maturity within the differentiated hepatocytes, comparative to the morphological data (Fletcher et al., 2008) (Hay et al., 2007) which was comparable to CM (Figure 3.9). For primer details please refer to table 2.4.

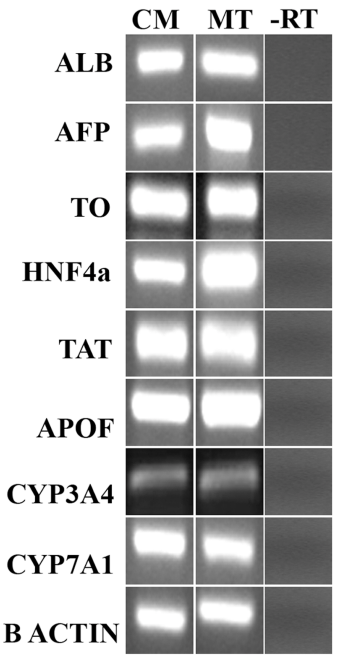


Figure 3.9 – Characterisation of HE derived from hESCs cultured in CM and MT.
Gene expression profiling demonstrates hepatic gene expression in HE derived from CM and MT cultured cells, which was carried out using RT PCR. HE derived from hESC cultured in both medias show comparable levels of hepatic gene expression, using a –ve RT as a control. Abbreviations: AFP – alpha fetoprotein; ALB – albumin; HNF4α – hepatocyte nuclear factor 4 alpha; CYP3A4 – cytochrome p450 3A4; TAT – tyrosine amino transferase; TO – tryptophan oxidase; APOF – apolipoprotein F; CYP7A1 – cytochrome p450 7A1

At day 17 of differentiation the purity of hESC-derived HE was quantified using an adult liver marker, albumin, and an IgG antibody was used as a negative control. hESC-derived HE purity was determined by the percentage of albumin-positive cells in differentiating cultures and estimated using four random fields of view with 500 cells per field of view. This data was further supported by estimating the percentage

of HE generated in 20 random fields of view per well based on morphology using phase contrast microscopy (n=3). We did not observe any significant differences regarding HE generation, in vitro, between each of the different media (Figure 3.10, graphs A and B). The CM control formed 96.5% (+/-1.5) HE and hESCs cultured in MT resulted in 98.6% HE (+/-0.38).

In addition to PCR, we confirmed hepatic gene expression by immunostaining fixed hESC-derived HE monolayers for a number of proteins involved in hepatic maintenance and function, (Section 2.3.1) Figure 3.10 A and B. hESC-derived HE maintained in CM and MT were fixed on day 17, and were stained with antibodies to albumin (ALB), alpha-feto protein (AFP), cytochrome p450 3A4 (Cyp 3A4), E-Cadherin, and hepatocyte nuclear factor 4 α (HNF4 α). Figure 3.10 displays the positive staining of the hepatic markers in HE derived from hESCs culture in CM (A) and MT (B). For antibody details please refer to table 2.3. Immunostaining was controlled using an IgG negative control.

The data presented demonstrates that hESCs cultured in MT can undergo efficient spontaneous and directed differentiation to human HE in vitro in a manner similar to CM control lines.

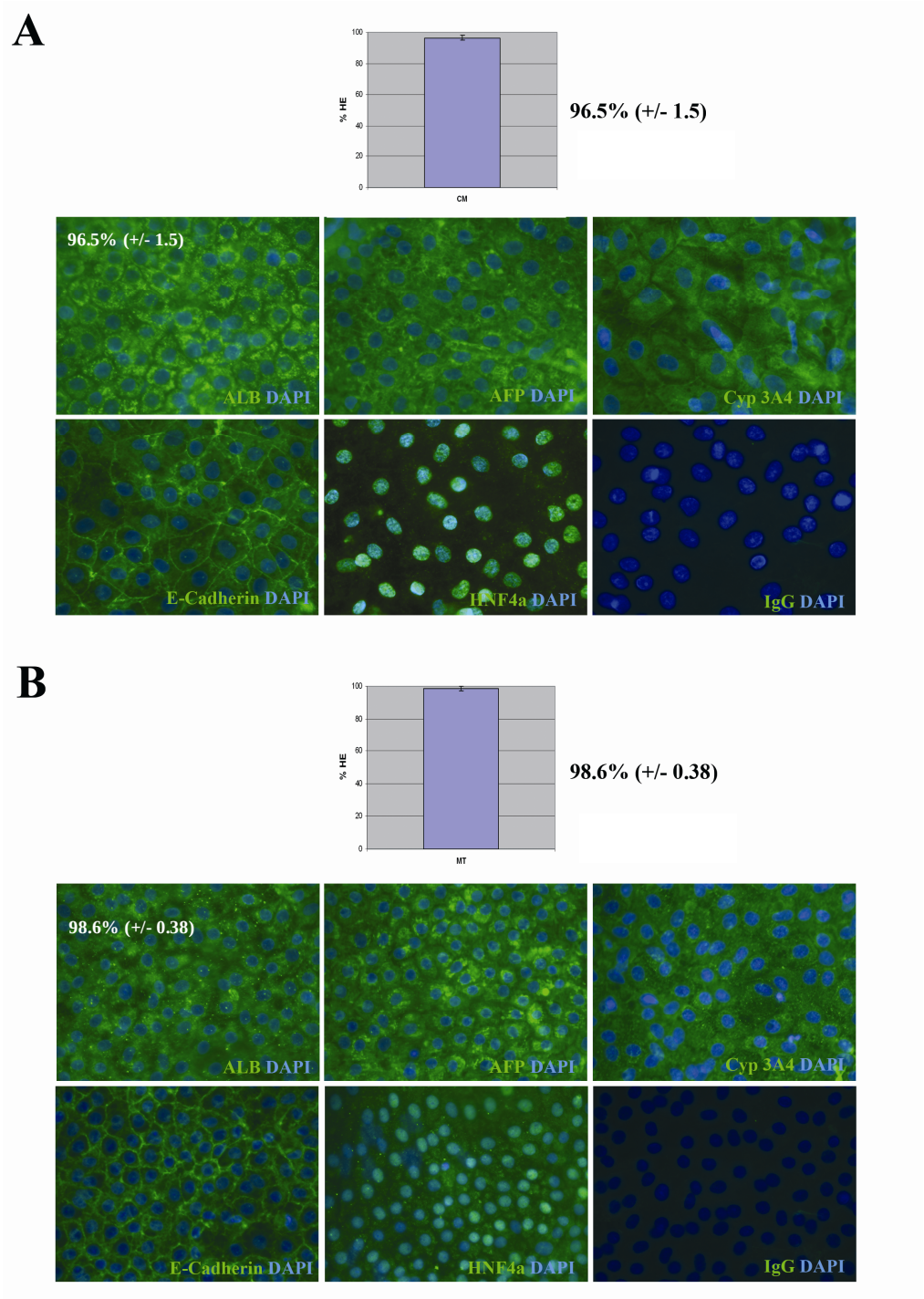


Figure 3.10 – Immunocytochemical Analysis of HE formed from hESCs cultured in CM and MT.

The graphs demonstrate the percentage yield of hepatic endoderm (HE) derived from hESCs cultured in the various media. Each media generated similar levels of HE. hESCs cultured in MT generated (**B**) 98.6% (+/- 0.38) HE and CM (**A**) 96.5% (+/- 1.5) HE. Four fields of view were counted, 500 cells in each view, and an additional 20 fields of view of phase contrast images were measured (n=3). IgG was used as a negative control. The images depict HE generated from hESCs cultured in CM (**A**) and MT (**B**). hESC-derived HE (Day 17) was stained for various hepatocyte markers, including albumin (ALB), alpha-feto protein (AFP), Cytochrome p450 3A4 (Cyp 3A4), E- Cadherin (E-Cad), hepatocyte nuclear factor 4-α (HNF4a) and an IgG control. Immunostaining was recorded using a Leica DMIRB inverted microscope, all images were captured at a x40 magnification.

3.2.3.3 FUNCTIONAL CHARACTERISATION OF HESC DERIVED HE *IN VITRO*

The production of functional HE from hESCs is essential for a number of downstream applications ranging from in vitro models to the bio-artificial liver and cell-based therapies. Following hepatic differentiation, we focussed on characterising hepatic function that would have direct application to the generation of predictive toxicology tools and extra-corporeal support devices. As such, we investigated a number of important liver functions including cytochrome p450 function (CYP3A4 and CYP1A2) (Section 2.3.5), ureagenensis (Section 2.3.6) and serum protein production, analysed using ELISA (Section 2.3.4).

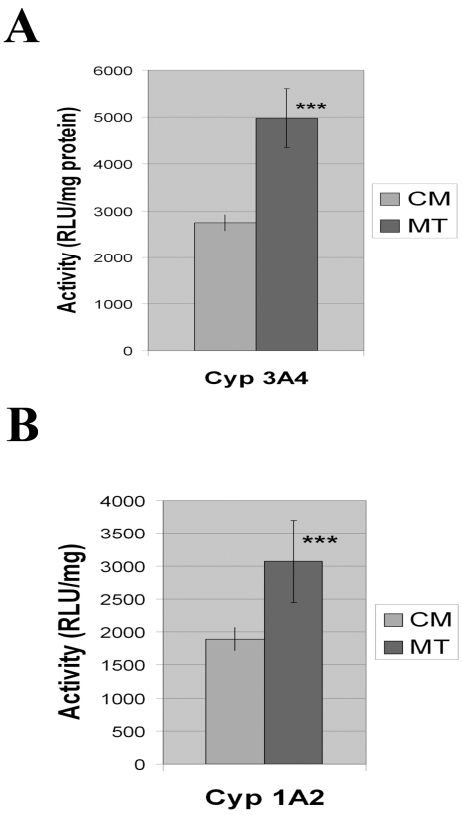


Figure 3.11 – Functional Analysis of HE derived from hESCs cultured in CM and MT, Cyp Activity.

hESC-derived HE cytochrome p450 function, Cyp3A4 and Cyp1A2 were assessed using the pGlo systems; (n=3) (Promega). **A.** Cyp3A4 activity, was approximately 2 fold greater than CM, observed in HE cultured in MT suggesting more mature hepatocytes with improved function , $p<0.002$ (ANOVA, $F=50.7$ and $R^2= 0.94$). **B.** hESC derived HE from MT cultures exhibited Cyp1A2 activity which was greater than that from CM cultures, $p<0.0001$ (ANOVA; $F = 64.68$ and $R^2= 0.96$). The values for Cyp3A4 and Cyp1A2 activity in CM and MT are as follows; 2646.7 RLU/mg (+/- 174) and 1822.9 RLU/mg (+/- 53), 4909.18 RLU/mg (+/- 630) and 3029.25 RLU/mg (+/- 259), respectively. The significance between differences in function of the various media was measured using ANOVA, and denoted on the graphs by an asterisk (*).

Both Cyp3A4 and 1A2 activity were observed in hESC-derived HE from cells maintained in CM and MT (Figure 3.11 A and B). Maximal CYP3A4 activity was detected in hESCs maintained in MT that was two fold higher than in CM-derived HE; ($p<0.0002$, ANOVA) (Figure 3.11 A). CYP1A2 activity was significantly greater in MT than that observed for CM; $p<0.0001$ (Figure 3.11 B). The values for Cyp3A4 and Cyp1A2 activity in CM and MT are as follows (+/- standard deviation); 2646.7 RLU/mg (+/- 174) and 1822.9 RLU/mg (+/- 53), 4909.18 RLU/mg (+/- 630) and 3029.25 RLU/mg (+/- 259), respectively.

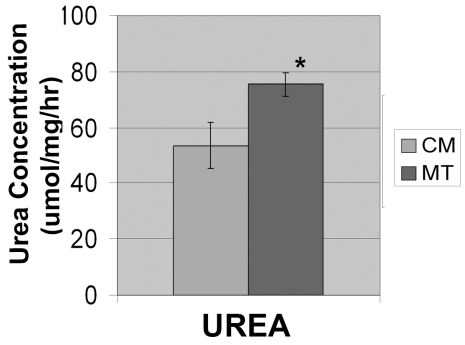


Figure 3.12 – Functional Analysis of HE derived from hESCs cultured in CM and MT, Urea Activity.
Urease function in HE derived from hESCs cultured in the various media is shown in the graph displayed. The level of ureagenesis was 1.5 higher in MT when compared to CM, $p=0.0118$, (ANOVA $F= 27.4$ and $R^2= 0.95$). This further suggests that MT derived HE is not only generated more efficiently but also shows increased function when compared to CM derived HE.

Urease activity was measured as previously carried out (Filippi et al., 2004). Urea production was 1.5 fold greater in MT than CM maintained hESCs; ($p<0.015$) (Figure 3.12). Urease activity in HE derived from hESCs cultured in CM and MT was measured to be 53.5 µmol/mg/hr (+/- 8.5) and 75.42 µmol/mg/hr (+/- 4.11), respectively, (+/- standard deviation).

Serum protein production is another vital function of mature hepatocytes required for healthy liver metabolism and homeostasis. Day 16 and 17 HE supernatants were collected and analysed for production and secretion of albumin, thyroxin binding pre albumin (TBPA) and fibrinogen; using ELISA.

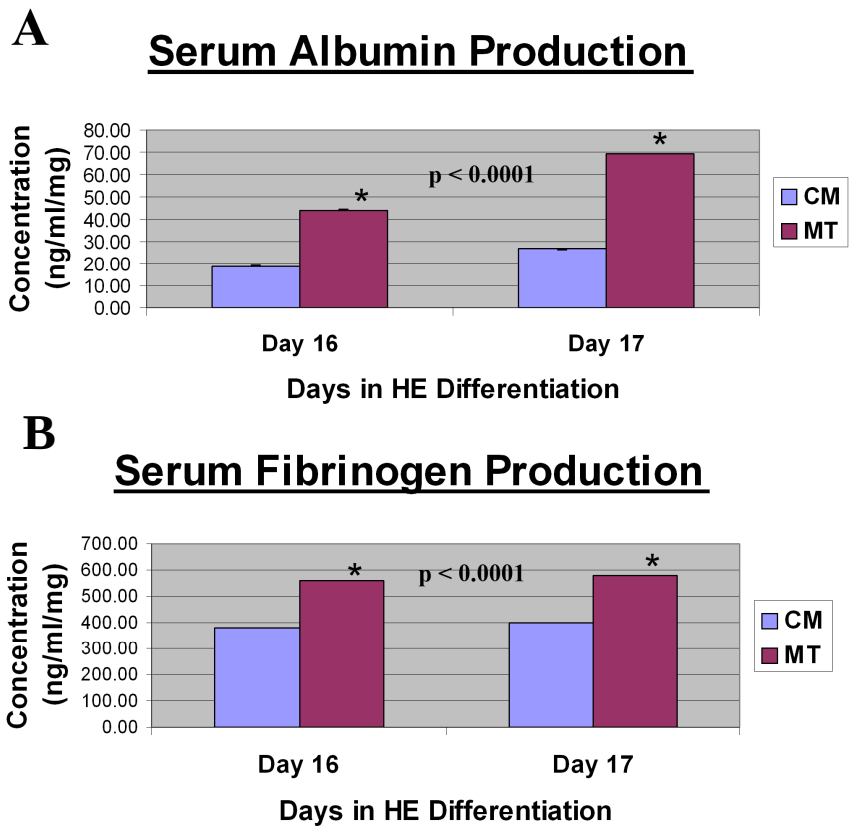


Figure 3.13 – Functional Analysis of hESC derived HE cultured in CM and MT, Serum Protein Production.

Serum protein production is another effective method to measure the maturity of hESC derived HE. The graphs displayed show the levels of Albumin (**A**) and fibrinogen (**B**) secreted by HE derived from hESCs cultures in CM and MT. These proteins are good indicators of hepatic maturity as they are uniquely produced by hepatocytes. As seen in both graphs, serum protein production for each protein is significantly higher in MT derived HE than CM, on both days 16 and 17 ($p<0.0001$). This defines MT derived HE as more functional than CM derived HE. The relative levels of serum protein production for CM derived HE are as follows; Albumin: D16 – 19 ng/ml/mg (+/- 0.29) and D17 - 26.91 ng/ml/mg (+/- 0.16), Fibrinogen: D16 – 377.86ng/ml/mg (+/- 0.6) and D17 – 401.18 ng/ml/mg (+/- 0.05). The relative levels of serum protein production for MT derived HE are as follows; Albumin: D16 – 44.33 ng/ml/mg (+/- 0.15) and D17 – 69.42 ng/ml/mg (+/- 0.12), Fibrinogen: D16 – 560 ng/ml/mg (+/- 0.06) and D17 – 581 ng/ml/mg (+/- 0.1).

Figure 3.13 demonstrates protein production for two proteins in HE derived from hESCs cultured in both CM and MT. The levels of albumin production increased significantly in both media from day 16 to 17, whilst fibrinogen levels were maintained over both days. The levels of albumin and fibrinogen were significantly higher in MT derived HE when compared to CM, ($p<0.0001$). The level of albumin

produced at day 17 was 2.6 fold higher in MT than CM and fibrinogen were 1.2 fold and 1.4 higher, respectively. This suggests that MT derived HE demonstrates increased function when compared to CM. The relative levels of serum protein production for CM derived HE are mentioned in Figure legend 3.13.

Sufficient levels of function was observed in HE derived from both media and were comparable to previous investigations. However, taken together, the data demonstrates that the introduction of a more defined cell culture technology not only permits scalable and reproducible hESC-derived HE, but also improves hepatic function when compared to CM-derived HE.

3.2.3.4 IN DEPTH PHENOTYPING OF HESC DERIVED HE

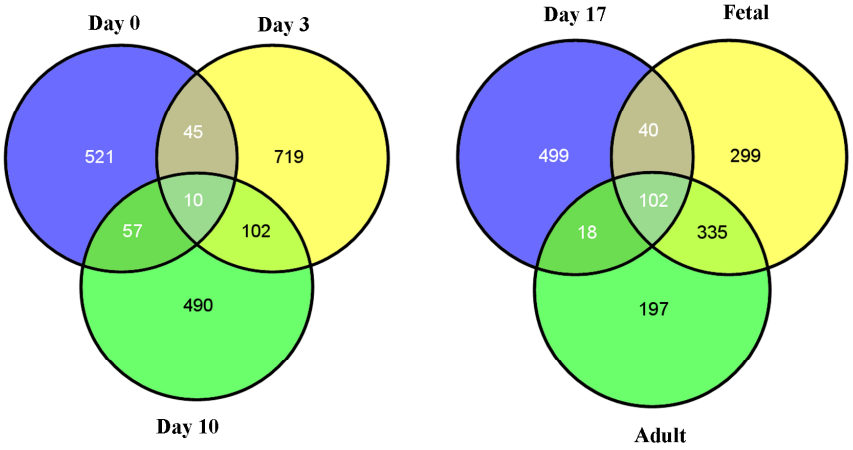
In order for hESC derived HE to be of use in an applied setting, they should acquire the majority of features associated with adult hepatocytes. As such, the genes expressed in hESC derived HE, fetal and adult hepatocytes were pooled and a comparative gene expression profile was generated. Initially, parameters were set to normalise and process the raw data, Figure 3.14 A. The total number of reads were measured followed by the total number of alignments, which generates a base number for further calculations. The aligned reads are then measured to signify a positive result, as it confirms that an mRNA strand were successfully aligned to a known gene. Unaligned reads are stands of mRNA sequences that could not be associated with any known gene, and this could be due to errors in the amplification stage or the integrity of the sequence itself. Discarded reads are not utilised as they are not an accurate representation of the mRNA.

A

	Day 0 - hESCs		Day 3 - hESC-ForegutEndoderm		Day 10 - Hepatoblasts	
#{reads}	19962278	---	29522918	---	22707888	---
#{total-alignments}	40356820	---	83173223	---	80382821	---
#{reads-aligned}	9464895	47,41%	19779458	67,00%	19042348	83,86%
#{reads-not-aligned}	10343004	51,81%	9448842	32,01%	3419809	15,06%
#{reads-discarded}	154379	0,77%	294618	1,00%	245731	1,08%

	Day 17 - Hepatocytes		Human-Fetal-Hepatocytes		Human-Adult-Hepatocytes	
#{reads}	12372693	---	29433136	---	30716279	---
#{total-alignments}	32865913	---	109169130	---	117552338	---
#{reads-aligned}	8253303	66,71%	25398556	86,29%	26477924	86,20%
#{reads-not-aligned}	4019962	32,49%	3761284	12,78%	3918252	12,76%
#{reads-discarded}	99428	0,80%	273296	0,93%	320103	1,04%

B



C

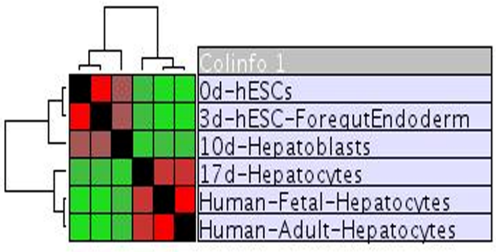


Figure 3.14 – Further Analysis of Gene Expression at Various Time Points.
A. The table displays a range of parameters and subsequent values used in defining gene expression at various time points throughout differentiation. **B.** The Venn diagram compares the genes expressed at the different time points. As observed; day 0, 3 and 10 have a few genes expressed in all three stages. However, when comparing day 17, fetal and adult samples, an increase in overlapping gene expression is observed. **C.** Sample clustering shows the relative levels of similarity in gene expression between the samples. Red defines close correlation whilst green defines distinct expression patterns.

Interestingly, there is a significantly higher number of non-aligned reads in the hESC sample. This could be due to the open chromatin structure usually associated with hESCs. As a result, RNA polymerase non specifically binds to DNA sequences, most likely non transcribed regions of the genome, and loosely expresses random DNA sequences.

The positive aligned reads each correspond to a gene, and are hence used to generate an expression pattern. The genes expressed in each of the samples were then compared, Figure 3.14 B and C. The Venn diagram, Figure 3.14 B, supports the differential gene expression patterns occurring at each specific time point. The samples were segregated into two groups; group one consisted of day 0, 3 and 10 samples whilst group 2 contained day 17, fetal and adult hepatocytes. As expected, only 0.5% gene expression similarity was detected between all three group 1 samples, attributed by the induction of a lineage change i.e.: converting hESCs into definitive endoderm, resulting in considerable alteration of the genes being activated. hESCs had approximately 3.6% and 4.6% gene expression similarity with the sample exhibiting primitive streak expression patterning and the sample expression genes observed in definitive endoderm formation, respectively. The comparative expression doubled to 7.3%, when comparing day 3 and 10 samples, which is due to the activation and maintenance of hepatic specification pathways required for differentiation in both samples.

The similarity in the expression pattern of all three group 2 samples increased by 10 fold when compared to group 1 samples, 4.9%, indicative of the constant expression

of signalling pathways vital for the maintenance of the mature hepatic phenotype. Interestingly, gene expression between day 17 hepatocytes was comparable to the fetal and adult hepatocyte samples, 9.9% and 9.1%, respectively, which rose to 30.1% when fetal and adult hepatocyte samples were compared. Therefore, Day 17 hepatocytes display comparable gene expression patterns to both the fetal and adult liver, contrary to the functional data previously noted in Section 3.2.3.3. It can be postulated that day 17 hepatocytes possess the correct gene expression pathways however, due to variations in translational processing the hepatic function is maintained at a more fetal than adult nature. In addition, the fetal expression patterns may be stalled in day 17 HE and an environmental cue is vital to drive maturation forward.

The data was also used to generate a heat map, Figure 3.14 C. The sample cluster shows that day 0, 3 and 10 samples show differential profiles of gene expression, displayed in red, which then changes as the cells differentiate towards hepatocytes suggesting similar gene expression patterns, marked by a colour change to green. As observed, day 17, fetal and adult hepatocytes cluster close together due to the similarity in their expression profiles. As the cells adopt a more hepatic nature, the majority of pluripotency genes are fully repressed and the expression of a large number of hepatic specific genes required for function and maturity are maintained; such as albumin, HNF4 α, HNF 6 and FOX A1 and thus the samples cluster together (day 17, adult and fetal). On the other hand, as hESCs differentiate into definitive endoderm, there is a significant modification in gene expression before the cells specialise down the hepatic lineage, post day 10, and therefore the samples exhibit

considerable difference in their gene expression patterns. Between days 0 and 10, the cellular phenotype is altered entirely, whereby genes required for pluripotency, Oct 4 and Nanog, begin to be repressed and genes required to initiate definitive endoderm differentiation are activated, GATA4 and FOXA1. Therefore, a differential gene expression pattern is generated between each sample (Day 0, 3 and 10) in conjugation with no similarity with the day 17, fetal and adult hepatocyte cluster.

Overall, the gene expression patterns within our hepatic differentiation model match those documented for human liver development, which supports its use as a reliable and accurate tool for investigation and as a robust starting point. Interestingly, the gene expression pattern of our hESC-derived hepatocytes is comparable to both fetal and adult gene expression suggesting the requirement of a switch to drive the complete maturation of day 17 HE. Therefore, modelling signalling pathways at a transcriptional and translational level will provide insight into the various methods of signalling regulation, whether transcriptional, translational or a combination of both.

3.3 DISCUSSION

Primary human hepatocytes are a limited resource with variable function that decreases with time in culture (Wege et al., 2003, Hay et al. 2007, Hay et al., 2008). Consequently, their use in therapeutic and clinical applications is restricted. As a result, other sources of hepatocytes have been explored to alleviate the dependence on primary hepatocytes. These sources include hepatocytes isolated from animal livers, hepatocarcinoma cell lines and derivates from adult stem cells (Medine et al., 2010). Unfortunately, each individual substitute has their disadvantages. The use of

animal hepatocytes, specifically porcine hepatocytes, in clinical applications may result in the risk of disease and xeno-genicity; as well as being an inaccurate representation when modelling human disease (Bonavita et al., 2010). Hepatocarcinoma cell lines can not be used for therapeutic or clinical purposes and have limited function in vitro (Iyer et al., 2010). The use of adult stem cells, such as hepatoblasts, is limited due to highly inefficient isolation and sub optimal purification. As a result, the derived hepatocytes are limited with unreliable function (Fiegel et al., 2006, Vessey et al., 2001, Alison et al., 2007).

Human embryonic stem cells (hESCs), on the other hand, have the potential to provide an inexhaustible supply of functional hepatic endoderm (HE) which, can then be used for applications such as drug discovery, disease modelling and cell therapy (Hannoun et al., 2010, Hay et al., 2008, Sharma et al., 2010). As a result, large amounts of time and resources, including our research group, have been invested into defining a standardised protocol for differentiating hESCs into functional and viable HE.

In vivo and *in vitro* formation of hepatic endoderm is a complex process regulated by growth factors, cytokines, transcription factors and the cellular adhesions. The essential factors include; FGF; a stimulator of hepatic gene expression and nascent hepatocyte stability *via* the activation of the RAS/MAP kinase pathway (MAPK) (Shin et al., 2007, Calmont et al., 2006). BMP signalling coincides with FGF signalling to sustain complete induction of hepatic endoderm. Co-operation between FGF and BMP 2 and 4, produced by the septum transversum mesenchyme, promotes

both competence and specification of the primitive endoderm down the hepatic lineage by induction of the transcription factor GATA4 (Huang et al., 2008). Wnt and β-Catenin are involved in both differentiation and proliferation of pre-hepatic endodermal cells (Burke et al., 2006, McLin et al., 2007, Fletcher et al., 2008). HGF binds to the c-Met receptor, activating both the SEK1/MKK4 and c-Jun signalling cascades resulting in Glucose-6-phosphate, Tyrosine amino transferase, Carbamoyl-phosphate synthase and Albumin expression, all of which are associated with mature liver (Duncan, 2003, Fiegel et al., 2008, Zaret, 2001). FOXA1, also known as hepatocyte nuclear factor HNF3α, is responsible for activating the expression of a large number of hepatocyte specific genes (Cereghini, 1996, Cirillo et al., 2002, Costa et al., 2003, Kaestner, 2000). The hepatocyte nuclear factors 1, 4 and 6 play a vital role in liver development (Odom et al., 2004, Lemaigre et al., 2004, Lokmane et al., 2008). Each factor has a specific function in hepatic specification, differentiation, maintenance and metabolic activity.

Elucidating the key pathways involved in hepatic endoderm formation *in vivo* has allowed us to apply these to the *in vitro* situation. There are a large number of protocols for HE formation ranging from embryoid body (EB) derivation in fetal bovine serum (FBS) (Rambhatla et al., 2003), to differentiating hESCs utilising 2-D systems using collagen, gelatin or Matrigel™ as the basement membrane. Agarwal and colleague's cultured hESCs on collagen in the presence of FBS, knockout serum replacement (KOSR) and bovine serum albumin (BSA) supplemented with Activin A (AA), fibroblast growth factor (FGF), hepatocyte growth factor (HGF), oncostatin M (OSM) and dexamethasone. The resulting HE expressed a number of hepatocyte

specific markers including albumin, alpha-fetoprotein (AFP), CYP3A4, CYP7A1 and was also capable of glycogen storage and albumin secretion (Agarwal et al., 2008, Hay et al., 2008). Schwartz and colleagues generated HE solely in the presence of FGF and HGF on collagen, expressing GATA4 and HNF1α and 3β; secreted human serum albumin, processed ammonia to urea production and exhibited Cytochrome P450 activity after treatment with Phenobarbital (Schwarts, 2005). HE has also been formed on collagen in the presence of FBS supplemented with insulin, dexamethasone, transferrin and selenious acid, where expression of albumin, transthyretin and albumin serum protein production was observed (Shirahashi et al., 2004). Hay and colleagues demonstrated that treatment with either Activin A/Wnt or sodium butyrate followed by treatment with dimethylsulfoxide (DMSO) generated immature hepatocytes. These hepatocytes could be matured with HGF, insulin, OSM and hydrocortisone to generate HE that expressed the majority of hepatocyte specific genes, were capable of glycogen storage, produced significant levels of hepatic serum proteins and had inducible Cyp P450 activity (Hay et al., 2008, Fletcher et al., 2008, Hay et al., 2007). These models all utilize factors associated with HE formation; however, each method differs in the factors used. Another difference between these systems is the varied use of the extra-cellular matrices. This highlights the importance of standardizing the differentiation process to allow for accurate reproducibility and modelling.

Culture of HE in 3-D environments will undoubtedly enhance HE function; as this potentially mimics *in vivo* development more accurately. It has been proposed that culturing hepatocytes between double layers of ECM in 3-D structures will establish

polarity and enhance hepatic function and *via*bility mimicking the *in vivo* situation. One group has successfully differentiated hESCs into HE in a 3-D environment. Baharvand *et al.,* cultured hESCs in self renewing conditions and using the hanging drop method formed EBs. These EBs were then seeded onto collagen coated 3-D scaffold in culture medium supplemented with FGF, HGF, OSM, insulin, dexamethasone, transferrin and selenium. The HE expressed a number of hepatic specific genes and produced significant levels of both urea and albumin (Baharvand et al., 2006). Du and colleagues successfully constructed an ECM free synthetic culture by sandwiching a hepatocyte monolayer between two membrane-like structures. The top support system consisted of a glycine - arginine - aspartic acid - serine (GRGDS) modified polyethylene terephthalate (PET) membrane. The bottom substratum consisted of a galactosylated PET membrane. This resulted in hepatic polarity including biliary excretion and enhanced function when compared to 2-D collagen coated cultures using HE derived from hESCs (Du et al., 2008). Ng and colleagues successfully combined methylated and galactosylated collagen nano-fibres that optimized the interactions required for the maintenance of functional hepatocytes. This enhanced interactions between the nano-fibres and the asialoglycoprotein receptor (ASGPR), hence promoting hepatic function (Ng et al., 2005, Yin et al., 2003, Lu et al., 2003). Hay and colleagues screened 380 polyurethanes and polyacrylates in order to identify a matrix that supported and promoted HE function and viability. Polyurethane 134 was found to not only enhance HE function but also permitted drug dependant induction of Cyp 3A4 with increased sensitivity over fresh isolated primary human hepatocytes (Hay et al., 2010). These results signify the importance of ECM interactions for maintaining hepatic functionality. Sufficient

differentiation has been achieved on 2-D culture systems. However, one can speculate that HE differentiation will never reach its full potential until culture systems incorporate the correct signalling factors and ECM.

Basma and colleagues established an efficient protocol for the purification of a hepatocyte population from a heterogeneous endodermal population. The hESCs derived EBs were plated onto Matrigel™ and treated with Activin A and FGF 2. The cells were then placed into defined media supplemented with HGF followed by dexamethasone. The resulting HE was further enriched by FACS sorting for ASGPR positive cells, a specific feature of mature hepatocytes. The pure population of HE expressed hepatic gene function comparable to adult hepatocytes (Basma et al., 2009).

Cai and colleagues developed a physiological protocol that mimicked the *in vivo* situation. This involves priming hESCs with Activin A to direct them towards definitive endoderm, followed by BMP and FGF generating hepatic endoderm. The resulting HE was matured using HGF, OSM and dexamethasone. The HE expressed a range of mature hepatic genes; however there was no expression of AFP, indicative that the HE produced is more mature than those derived using other protocols. The HE produced significant albumin and interestingly was susceptible to infection by the hepatitis virus (Cai et al., 2007).

As the field of regenerative medicine advances there will be a requirement for more defined and reproducible culture systems. This has been the goal of our lab, as

described in the above sections, and Baharvand and colleagues. hESCs have been grown on Matrigel™ under serum free conditions and treated with a series of defined factors to produce HE. The resulting HE expresses hepatic genes, produce serum protein, display, urease activity, glycogen storage and uptake of low density of lipoproteins (Baharvand et al., 2008, Hannoun et al., 2010).

A combination of studies, including our own, has demonstrated that MT is a suitable and more defined media than CM for culturing hESCs. hESCs maintained in MT retained all embryonic stem cell specific phenotypes, Figures 3.1 – 3.4, and in our system demonstrated an increased capacity to form hepatocytes with improved function, Figures 3.7 - 3.13. In conjunction with the functional data, we observed the expected patterns in gene expression when compared to liver development in vivo, Figure 3.7. This model provided a starting point to further elucidate the important pathways required for efficient hepatocyte differentiation.

A large number of factors affect the generation, function and viability of hepatic endodermal cells. As a result, isolating a few vital elements to fine tune our system has proven difficult. However, preliminary data in conjunction with a thorough literature review has deemed hepatocyte nuclear factor 4 α (HNF4 α) as a possible candidate. The next chapter focuses on HNF4 α, an essential factor required for both liver development and metabolism, and its role and subsequent use in refining our differentiation model, which in turn can be translated for application in other systems.

CHAPTER FOUR

THE IMPORTANCE OF HEPATOCYTE NUCLEAR FACTOR 4 α IN HEPATIC ENDODERM DIFFERENTIATION

CHAPTER 4

THE IMPORTANCE OF HEPATOCYTE NUCLEAR FACTOR 4 α IN HEPATIC ENDODERM DIFFERENTIATION

4.1 INTRODUCTION

4.1.1 PROPERTIES AND CHARACTERISTICS OF HNF4 α

Hepatocyte nuclear factor 4 α is a member of the nuclear hormone receptor family (Watt et al., 2003) and is highly conserved (Sladek et al., 1990). It is known as an orphan receptor as there is no known ligand and has been defined as constitutively active (Sladek et al., 1990, Dhe-Paganon et al., 2002). The protein is predominately expressed in the liver, pancreas, kidney, colon and intestine. HNF4 α induces gene expression by binding to a 6 base pair repeat (AGGTCA), known as direct repeat 1 (DR1) located upstream to its target genes and binds strictly as a homodimer (Harish et al., 2001, Peiler et al., 2000). Unlike other nuclear receptors, it cannot form heterodimers (Jiang et al., 1997, Bogan et al., 2000). Studies have shown that HNF4 α is essential during liver development (Hayhurst et al., 2008, Ryffel, 2001). Knocking out HNF4 α during development inhibits liver formation and severely disrupts extra embryonic endoderm generation (Chen et al., 1994, Peiler et al., 2000). HNF4 α expression is also crucial in the adult liver; controlling the expression of a large number of genes, approximately 50% of the hepatic transcripts, required for normal metabolic functions (Li et al., 2000). HNF4 α regulates the expression of genes associated with functions such as; urea production, drug metabolism, bile acid

synthesis, fat metabolism and glucose homeostasis (Li et al., 2000, Watt et al., 2003, Wortham et al., 2007, Ellard S FAU - Colclough et al.,))

4.1.2 THE STRUCTURE OF HNF4 α

HNF4 α possesses a molecular structure common to other transcription factors, some of which are nuclear receptors. Amino acid residues 1-50 contain an activation domain responsible for transcriptional activation (AD-1), followed by the DNA binding (DNA BD) domain which is conserved in most nuclear receptors, located between residues 50-140, (Figure 4.1 A). The DNA BD consists of two zinc fingers joined by a hinge region to allow conformational flexibility required for function. The ligand binding domain (LBD) consists of alpha helix 9 and 10, spanning residues 141-368, and is proceeded by another activation domain (AD2) (Figure 4.1 A). Residues 369 to 465 contain the repression domain, of which its effect may be circumstantially alleviated (Sladek et al., 1999, Hadzopoulou-Cladaras et al., 1997). The LBD is important for transcriptional activity and has two distinct conformations, open and closed, differing by the position of helix α12 (the AF-2 domain) (Figure 4.1 B). Helix α12 is extended co-linearly with helices α10 and α11 in the open configuration. This renders the ligand binding pocket accessible, however, this open structure prevents the binding of co-activator molecules; thus HNF4 α is in an inactive form. The α12 helix folds back onto the LBD in the closed conformation, restricting access to the ligand binding pocket. The closed state of HNF4 α allows interactions with co-activators such as C/EBP, creating the HNF4 α active state. Duda et al has provided evidence that although HNF4 α ligand binding and co-activator binding both stabilize the LBD, it is co-activator binding that locks HNF4 α

in the active state. HNF4 α has also been shown to reversibly bind to a single fatty acid, linoleic acid, suggesting the potential use of HNF4 α as a drug target for various liver disorders. (Bogan et al., 2000, Dhe-Paganon et al., 2002, Duda et al., 2004, Ryffel, 2001, Sladek et al., 1990)

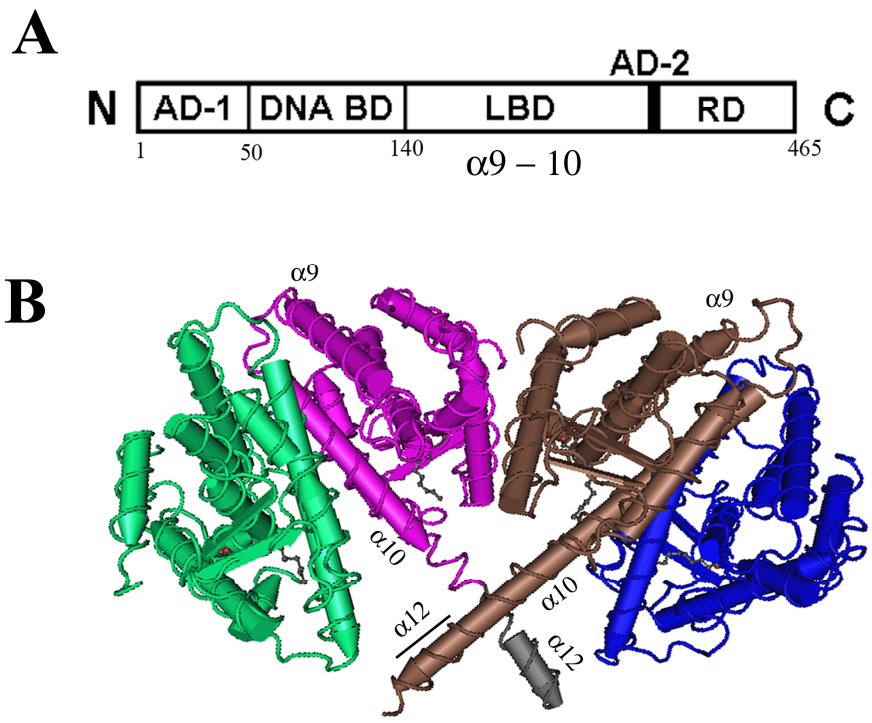


Figure 4.1 – Structure and 3D Schematic of Hepatocyte Nuclear Factor 4 α.
A. The diagram shows the functional groups of HNF4 α, including the activation domain (AD), the ligand binding domain (LBD), the DNA binding domain (DNA BD) and the repression domain (RD). (Bogan et al., 2000, Sladek et al., 1990) **B.** The 3D schematic of the HNF4 α heterodimer shows the subunits in the open conformation allowing for ligand binding, as helix twelve is aligned with helix 10 leaving the binding domain accessible. (Dhe-Paganon et al., 2002)

4.1.3 HNF4 α Function in the Cell

HNF4 α has been implicated in processes essential for development, hepatic differentiation and normal adult liver metabolism (Sladek et al., 1999). HNF4 α was found to bind to 80% of the genes expressed in hepatocytes, when analysed using the Hu13K Microarray, with 42% of these genes also being occupied by RNA polymerase II, confirming its crucial role in hepatocytes (Odom et al., 2004). Studies

show that embryos deficient in HNF4 α die during gastrulation (Chen et al., 1994). This can be circumvented using conditional knockouts, whereby mice containing lox p sites flanking exon 2 of the HNF4 α gene were mated with mice expressing Cre recombinase in the fetal liver under the control of the albumin promoter and the AFP enhancer region. Heterozygous double breeding resulted in the formation of embryos containing the lox p flanked HNF4 α gene in addition to the Cre recombinase gene regulated by albumin and AFP expression. However, this resulted in the formation of embryonic livers containing large red lesions, with discontinuous parenchyma (Parviz et al., 2003). In addition, HNF4 α directly activates the expression of hundreds of genes; especially those prevailing in processes such as drug metabolism and glucose and fatty acid regulation (Odom et al., 2004). The deletion of HNF4 α has deleterious effects for hepatocyte differentiation, including metabolic function and altered cell morphology and adhesions (Parviz et al., 2003). HNF4 α has also been associated with healthy pancreatic function, indicative of its importance in regulating endodermal differentiation as opposed to just hepatic. Mutations in the HNF4 α gene in the pancreas results in Maturity Onset Diabetes of the Young (MODY1) (Yamagata et al., 1996). This disease is a form of type 2 diabetes, characterised by high levels of glucose in the blood due to insulin deficiency and resistance (Yamagata et al., 1996). Another effect of an HNF4 α mutation includes increased body weight, hyperglycaemia and hyperinsulinemia in HNF4 α heterozygous children at birth (Lemaigre et al., 2004). Additionally, various mutations in the HNF4 α gene have been detected in patients suffering from diabetes (Stanger, 2008).

4.1.4 REGULATING HNF4 α FUNCTION

HNF4 α is known to be constitutively active and found in the nucleus at a steady state; however, the regulatory mechanisms which govern HNF4 α activity are poorly understood. Nonetheless, data has depicted the importance of HNF4 α phosphorylation in transcriptional activation (Sun et al., 2007). HNF4 α contains a conserved serine residue, S78, which is phosphorylated by protein kinase C (PKC), in vitro (Sun et al., 2007). Generation of a phosphomimetic mutant leads to reduced DNA binding, decreased transcriptional activity, re-localisation to the cytoplasm and enhanced proteasome mediated degradation of the HNF4 α (Sun et al., 2007). Similar observations have been noted when HNF4 α is phosphorylated by protein kinase A (Viollet et al., 1997), AMP activated protein kinase resulting in reduced dimer formation (Hong et al., 2003) and p38 a vital factor in the inflammatory response (Guo et al., 2006). Various HNF4 α co-activators have been established such as; p300, CBP, GRIP 1 and SRC1a (Sladek et al., 1999). Another study correlated phosphorylation of another serine residue, S158, in response to oxidative stress and IL-β1 resulting in the enhancement of transcriptional activity (Guo et al., 2006). Various investigations have identified fatty acids as a target ligand for HNF4 α and in turn affect its activity (Dhe-Paganon et al., 2002, Wisely et al., 2002, Hertz et al., 1998). Furthermore, Yuan and colleagues demonstrated the reversible binding of the fatty acid, linoleic acid, with no effect on transcriptional activity (Yuan et al., 2009). Despite the above information, the complete mechanisms behind HNF4 α regulation and function are yet to be elucidated, more specifically its role in hepatic differentiation and the maintenance of HE viability, which has become the focus of this thesis.

4.1.5 HNF4 α SPLICE VARIANTS

The presence of two HNF4 α promoters within the gene results in the expression of six known splice variants (Harries et al., 2009, Drewes et al., 1996). Variants 1, 2 and 3 are usually expressed in the liver, whilst variants 4, 5 and 6 (also referred to as 7, 8 and 9) are found in the pancreas, figure 4.2 (Harries et al., 2009). As depicted in figure 4.2, HNF4 α variants 1, 2 and 3 rely on promoter P1 for expression, differing from v 4, 5 and 6 that utilise P2, also present in the fetal liver, established in murine development (Harries et al., 2008, Drewes et al., 1996, Eeckhoute et al., 2003). The presence of two promoter regions also allows spatial and time dependant induction of HNF4 α expression. The expression of both P1 and P2 transcripts occur during development; however, P1 transcripts are expressed exclusively in the adult liver with significant down regulation of P2 driven transcripts, in the mouse (Stoffel et al., 1997) (Briancon et al., 2004). The switch between HNF4 α promoters has been suggested to be key in achieving a mature hepatic phenotype. The expression of P1 promoter driven transcripts are dependant on HNF 1α and 1β, GATA-6, HNF 6 and RXR/RAR. In addition, the HNF4 α enhancer region has been found to be bound by HNF 3, HNF 1α and 1β, HNF4 α and C/EBP α (Stoffel et al., 1997). On the other hand, studies have shown that HNF4 α v1 is responsible for repressing the transcriptional activation of the P2 driven variants within the adult liver (Briancon et al., 2004). The mechanism proposed relies on negative feedback regulation whereby the initial activation of P1 HNF4 α variants in the fetal liver results in the repression of the P2 variants in the adult liver, indicative of the importance of the promoter switch. In addition, HNF4 α P1 variants promote self-transcription by binding to the P1 enhancer region, resulting in sustained expression of HNF4 α required in the adult

liver (Ribeiro et al., 2007). Interestingly, there is no published data specifying the variant expression pattern observed in the human fetal and adult liver.

The structure of HNF4 α has been demonstrated to play a key role in regulating its functional activity. The LBD and DNA BD are conserved between all six variants; however variants differ in their activation domain (AD) and their repression domains (RD) (Figure 4.2) (Drewes et al., 1996, Sladek et al., 1990, Duda et al., 2004, Sladek et al., 1999). This difference may be responsible for altering the repertoire of genes activated within

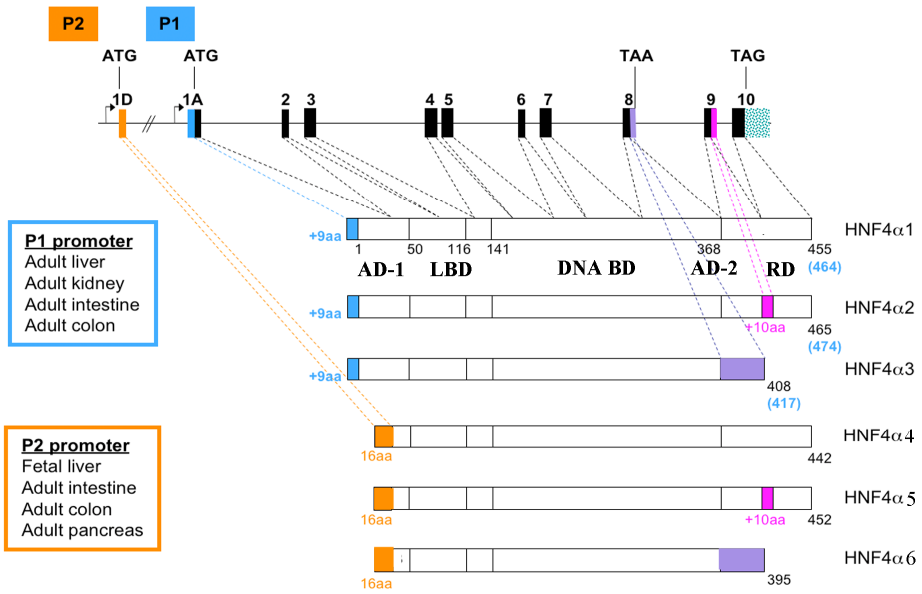


Figure 4.2 – The Different HNF4 α Splice Variants.
The diagram above shows the changes in the HNF4 α structure between the six different splice variants. All variants contain the ligand and DNA binding domains and differ in their activation and repression domains. Variants 1, 2 and 3 rely on promoter one for expression, used in the adult liver whilst variants 4, 5 and 6 rely on promoter two, used in the fetal liver. (Sladek et al., 1990)

the fetal liver, pancreas and adult liver. Studies indicate an anti tumorigenic role for P1 driven HNF4 α variants; moreover, over expression of these variants reduces the

proliferation of renal cells demonstrated in HEK293 cells (Lucas et al., 2005) as well as hepatocellular carcinomas observed in the murine embryo tumorigenic cell line, F9 (Chiba et al., 2005). On the other hand, high levels of the P2 driven variants were detected in cancerous regions (Tanaka et al., 2006). Variant five, a P2 promoter driven product, displays limited transactivation capabilities when compared to variant 2, this may be due to differences in the N terminal activation region and the C terminal repressor domain (Ihara et al., 2005). When comparing variants 1 and 4, the only notable splicing difference is the lack of the AD-1 domain in variant 4, which results in a dramatic change in function, whereby specific co-activators are not able to enhance HNF4 α functional activity (Torres-Padilla et al., 2002). The collection of this data suggests that each variant may have a specific function, differing not only in their target genes but also in tissue specificity. This emphasises the lack of redundancy within the HNF4 α signalling pathway, further supporting the importance of HNF4 α as a key factor in hepatic differentiation, postnatal liver function and tissue homeostasis.

4.1.6 THE HNF 1, HNF 4 α AND HNF 6 SIGNALLING NETWORKS

Signalling networks in hepatocytes and pancreatic cells have been thoroughly investigated. Three key factors play major roles in both the development and healthy function of the liver and pancreas; HNF4 α, HNF1 α and HNF 6. For the purpose of this thesis, I will only focus on their roles in the liver. As previously mentioned, HNF4 α is part of the nuclear receptor superfamily that strictly binds DNA as a homo-dimer. HNF1 α is a POU homeodomain protein that binds DNA as a homo and hetero-dimer (Costa et al., 2003) and binds to genes involved in lipid metabolism,

serum protein production and detoxification. HNF6 is a ONECUT transcription factor where it binds DNA with its cut-homeodomain as a monomer (Costa et al., 2003) and alters the expression of genes vital for cell cycle regulation and acts as a master regulator for feed-forward mechanisms in the liver. All three-transcription factors contain DNA binding domains; however, they differ significantly in their structure, although each DNA BD is conserved within each protein. HNF1 α dimers form a unique four helix domain required for DNA binding and co-activator interaction (Rose et al., 2000). HNF 6, on the other hand, binds DNA as a single protein via the cut domain (Sheng et al., 2004). HNF 6 also contains another DNA binding homeodomain; however, it can only bind DNA if the cut region is already bound (Sheng et al., 2004). HNF4 α, like HNF1 α, can only bind DNA as a dimer via two zinc fingers. Both HNF1 α and 4 α also contain transactivation domains (Navas et al., 1999).

HNF4 α and HNF1 α bind to each others promoter regions, creating a multicompentent loop, (Figure 4.3 B) (Odom et al., 2004). This mechanism allows for feedback control and the flexibility within the system to switch between the two different states. HNF1 α, 4 α and 6 also control HNF4 α gene expression (Figure 4.3 A). However, it has also been observed that HNF4 α and HNF6 interact and cooperatively bind promoter regions of genes required for gluconeogenesis (Odom et al., 2004). Another set of genes require all three-transcription factors to be bound to its promoter region for expression, whereby inhibiting the binding of one factor inhibits its transcription (Odom et al., 2004). These genes include histamine N-methyltransferase and the alpha-1-microglobulin/bikunin precursor both of which are

involved in the inflammatory response. The complex networking between these transcription factors allows for intricate regulation of the processes required for healthy hepatic function.

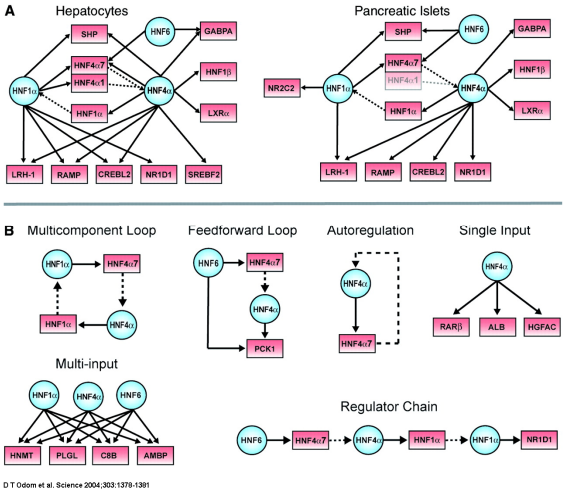


Figure 4.3 – The Transcriptional Networking Between HNF 1 α, 4 α and 6.
Diagram **A** shows the transcriptional network of HNF1α, HNF6, and HNF4α. The blue circles depict target genes and the red boxes represent their target proteins. Solid arrows show DNA- protein interactions, whilst dashed lines indicate proteins regulated by the transcriptional products of the genes. Overall, these genes regulate a large number of genes and do so in a timed and cooperative manner. Diagram **B** displays examples of regulatory network motifs in hepatocytes taken from (Odom et al., 2004).

4.2 RESULTS

Given the key role of HNF4 α in hepatic differentiation, hepatocyte function and viability, it was vital to elucidate the exact mechanisms behind the regulation and function of HNF4 α within our in vitro model. Furthermore, regulating the levels of HNF4 α within our system may enhance HE function and viability.

4.2.1 HNF4 α PATTERNING IN LIVER DIFFERENTIATION

hESCs cultured in both CM and MT were differentiated using standardised protocols (Hay et al., 2008) and cell extracts were collected at key time points throughout the differentiation process. The cells were lysed with a SUMO specific lysis buffer (Table 2.1) containing iodoacetamide, a compound that specifically inhibits SUMO proteases. The cell extracts were then separated using SDS PAGE and probed with specific antibodies targeting key proteins for each time point using standard western blotting. The hepatocyte nuclear factor 4 α (HNF4 α) antibody was used to detect HE specification (Figure 4.4). Similar patterns were observed in both hESC derived HE cultured in CM, (Figure 4.4 A) and MT (Figure 4.4 B).

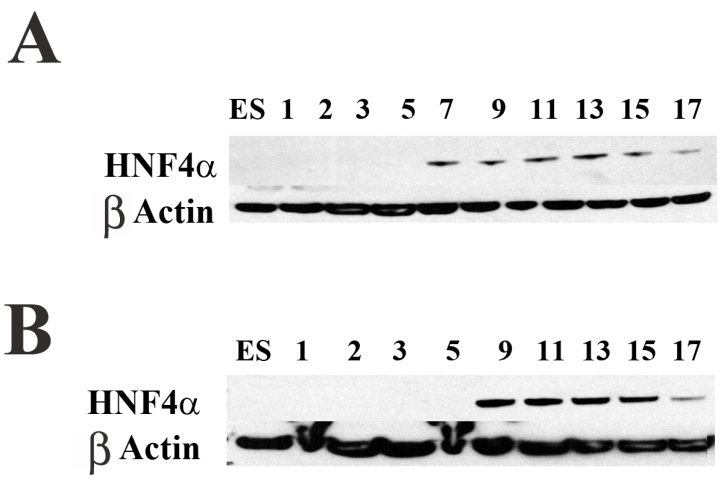


Figure 4.4 – Investigating HNF4 α in HE derived from hESCs Cultured in CM and MT.
Cell extracts were collected at various time points throughout the differentiation protocol (Hay et al., 2008) and were probed with hepatocyte nuclear factor 4 α (HNF4α) antibody using standard western blotting. HNF4α depicts hepatocyte specification. A decrease in the levels of HNF4α was observed as the hepatocytes mature. Similar patterns are observed in both medias, however, HNF4 α expression is induced at day 7 in CM (**A**) and day 9 in MT (**B**). Panel **A** shows extracts from hESC derived HE cultured CM and panel **B** displays hESC derived HE cultured in MT. β actin was used as a loading control.

Peak levels of HNF4 α, an important transcription factor in both the developing and adult liver, were noted at day 7 in CM and day 9 in MT; the stage of hepatic

commitment. High levels of HNF4 α have also been documented within the developing liver at the point of hepatic specification and commitment. In conjunction with this observation, it was noticed that the levels of HNF4 α decreased to levels required to activate mature factor expression, as noted in the developing liver. Interestingly, the secreted levels of thyroxin binding pre-albumin (TBPA), increased by more than 2-fold, and fibrinogen, both HNF4 α targets, increased by less than 2-fold between days 15 and 17 coinciding with the decrease in HNF4 α (Figure 4.5) suggesting a possible role of HNF4 α in regulating metabolic functions in our differentiation system. However, levels of HNF4 α were barely detectable at day 17, which suggests that HNF4 α may play a significant role in HE de-differentiation and decreased cell viability.

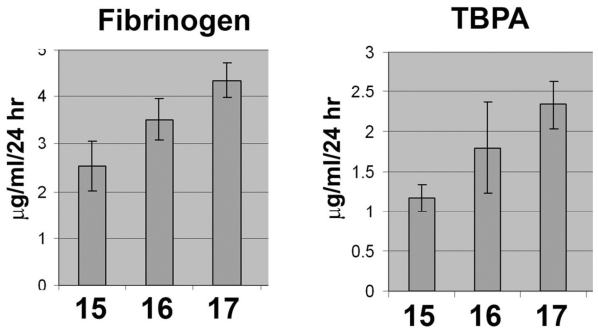


Figure 4.5 – Investigating the Effect of HNF4 α on HE Function.
Serum protein production of fibrinogen and thyroxin binding pre albumin (TBPA), known targets of HNF4 α, were measured at day 15, 16 and 17. The levels fibrinogen increased by less than 2 fold and TBPA increased by more than 2-fold from day 15-17 coinciding with the decrease in HNF4 α. The serum protein concentrations were measured using ELISA. For details please refer to (Hay et al., 2008).

In conclusion, HNF4 α expression patterns within our differentiation protocol mimics changes noted throughout liver development (Lemaigre et al., 2004, Costa et al., 2003), illustrating the accuracy of our differentiation model. However, unlike the adult liver, the levels of HNF4 α at day 17 HE are notably low; which is also observed in day 5-7 old adult hepatocytes in vitro (data not shown), as the levels of

HNF4 α decrease with time in culture. During liver development, the levels of HNF4 α expression significantly decrease between the fetal and adult liver. Nonetheless, the levels never decrease to the point observed in day 17 HE. The considerably low levels of HNF4 α in day 17 HE and the in vitro cultured primary hepatocytes may account for their limited viability, in vitro. In theory, stabilising HNF4 α within our hepatocytes may prevent their de-differentiation and subsequent death in long-term in vitro culture. This hypothesis forms the basis of my thesis.

4.2.2 Investigating HNF4 α Splice Variant Expression in HE Differentiation

Section 4.1.5 highlights the importance of the different HNF4 α variants in regulating cellular phenotype, however limited knowledge is available in the expression pattern in the adult and fetal liver in humans. Consequently, the expression patterns of the HNF4 α variants within our system was examined to enhance our understanding of signalling with our model. This was carried out using reverse transcription (RT) PCR. Initially, positive controls for each variant were generated by designing variant specific primers. The primers were then used to amplify variant specific sequences from an HNF4 α image clone (IC:9021539) using the standard PCR technique (Figure 4.6 A), thus generating the positive controls. The products were then run on an agarose gel, purified using a gel extraction kit, ligated into the PCR 2.1 vector and subsequently amplified and purified using the QIAGEN mini prep kit. PCR optimisation was then carried out by altering the annealing temperature to improve primer

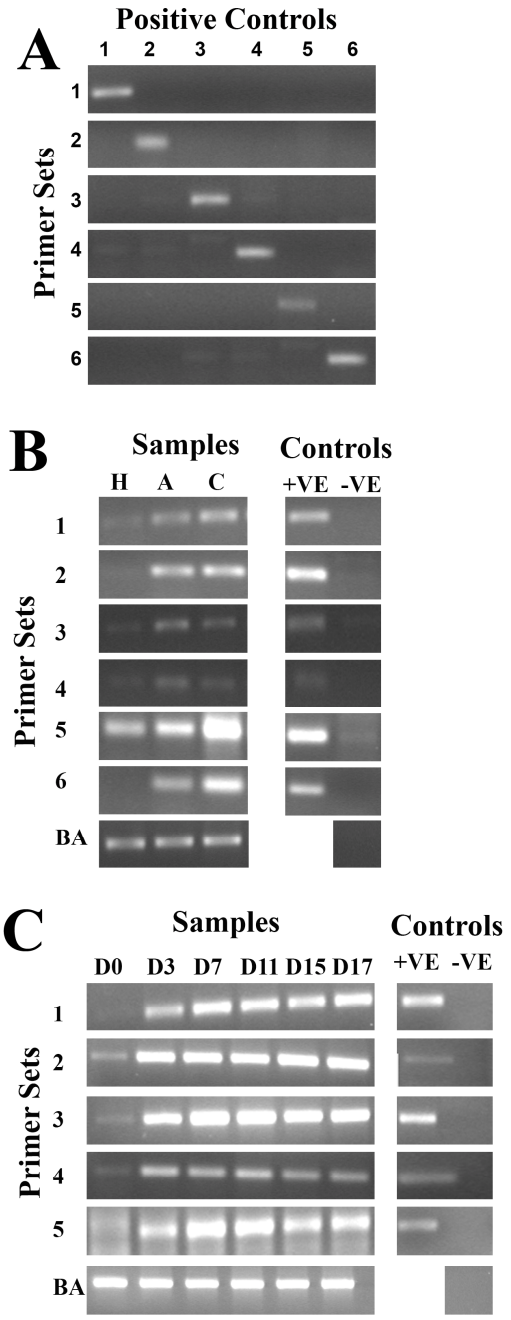


Figure 4.6 – Analysis of HNF4α Variants in Various Cell Types.
A. The PCR gel displayed shows the positive controls generated to test each specific primer set that should recognise the individual HNF4α variants. **B.** Panel B displays the HNF4 α variants present in human embryonic stem cells (H), adult hepatocytes (A) and day 17 HE (C). Variant 1, 2, 5 and 6 are expressed at equivalent levels in adult hepatocytes and in HE variants 5 and 6 seem to be slightly up regulated. A band present in primer set 5 in the hESC suggests a false positive. **C.** The gel displays the analysis of the HNF4 α variants throughout the differentiation protocol. All variants are initially expressed at day 3 and expression levels do not vary throughout the rest of the differentiation protocol. B-Actin (BA) was used as a loading control.

specificity in conjunction with cycle number to ensure the product was not at saturated levels. This allows direct comparisons between the samples. Once the PCR conditions were optimised, RNA was extracted from hESCs (H), adult liver (A) and hESC derived HE (C) and transcripts amplified using the variant specific primers. The products were run on a gel as well as the positive and negative controls (Figure 4.6 B).

The gel shows positive expression for all variants in the adult and HE samples and a slight expression of v3, 4 and 5 in hESCs. This could be due to spontaneous differentiation within our hESC cultures. RNA was also extracted from various time points throughout the HE differentiation and transcripts amplified as previously mentioned. The gel in figure 4.6 C shows expression of variants 1-5 from day 3 onwards and no real difference between the variants themselves. Variant 6 was unable to be amplified again in the samples and positive control despite repeated attempts and optimisation efforts, which may be due to primer annealing issues. Variants 4 and 5 are expressed in our HE, which could explain the fetal nature of the hepatocytes supporting previous data (Chapter 3). The switch between HNF4 α variants may be key in generating mature hepatocytes, as observed in mice (Sladek et al, 2000). However, due to the inaccuracy of RT-PCR for transcript quantification, no definitive conclusions were drawn. In order to quantify the expression of the HNF4 α variants, SOLEXA was employed. The data generated by SOLEXA allows absolute quantification and enhances the specificity of detecting each transcript.

4.2.3 ANALYSING HNF4 α SPLICE VARIANT EXPRESSION USING SOLEXA

SOLEXA analysis was used not only to identify the expression pattern of the HNF4 α variants within our differentiation model but to also identify the pattern between fetal and adult hepatocytes. hESCs were differentiated into HE using the standardised procedure and RNA was collected at day 0, 3, 10 and 17. The cells adopted the expected morphology of primitive endoderm, day 3; hepatic specification, day 10 and maturity, day 17 (Figure 4.7). Day 17 hepatocytes were positive for albumin expression, with 91% (as calculated in section 3.2.3) HE suggesting a pure population, a pre-requisite for accurate transcript quantification (Figure 4.8).

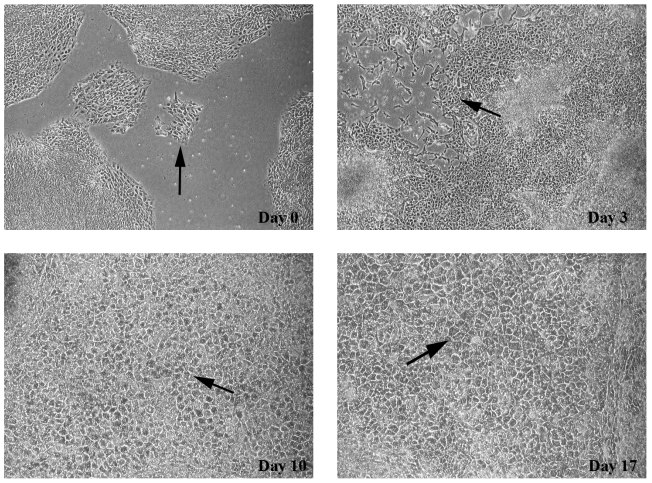


Figure 4.7 – Morphological HE Characterisation of Samples Generated for use in SOLEXA.
hESC were differentiated to HE and RNA samples were collected at day 0, 3, 10 and 17. The RNA was further purified and analysed using SOLEXA. This panel shows the morphological changes that occur throughout differentiation. At day 0, the ES colonies are well defined with uniform cell structure. At day 3 primitive endoderm like cells begin to form; this is followed by the formation of large hexagonal shaped cells at day 10. Day 17 HE show similar morphology of that observed in adult hepatocytes.

The RNA was extracted from samples collected at days 0, 3, 10 and 17, as well as fetal and adult hepatocytes and was isolated using the technique described in section 2.4.1 and sequenced using SOLEXA. The SOLEXA technology relies on the ability to amplify a single target of DNA molecules using PCR. The mRNA is converted to cDNA, which is then tagged to identify the various samples. The mRNA binds to the

chip and is amplified and sequenced to identify each transcript expressed within the sample. The values are normalised and subsequently analysed.

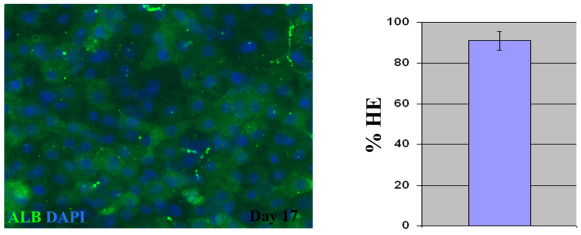


Figure 4.8 – HE Characterisation of Samples Generated for use in SOELXA.
The figure depicts HE generated from hESCs cultured in CM. hESC-derived HE (Day 17) was stained for a hepatocyte marker, albumin (ALB). Immunostaining was recorded using a Leica DMIRB inverted microscope; all images were captured at a x40 magnification. The graph shows 91% HE formation, four fields of view were counted, 500 cells in each view, and an additional 20 fields of view of phase contrast images were measured (n=3, 3 different wells). IgG was used as a negative control.

The pattern observed in HNF4 α expression of all the variants mimic what was noted in western blotting. The levels of HNF4 α increase to a peak level at day 10 (~20 RPKM, reads per Kilo base of exon per million mapped reads) followed by a significant decrease at day 17 (Figure 4.9). Despite the decrease of HNF4 α transcription at day 17 (~9 RPKM), the reduction in transcript levels does not account for the barely detectable levels of HNF4 α seen in the western blots. This data suggests that HNF4 α is not only regulated at a transcriptional level but also post-transcriptionally. There is a modest difference in the expression values of the various variants throughout the differentiation process, each averaging 10-20 RPKM (reads per Kilo base of exon per million mapped reads) (Figure 4.9). The lack of differential expression between the variants and the low expression of variant 3 within the day 17 HE may correlate to their subsequent arrest at the fetal stage. Unfortunately, in accordance with previous data, the levels of HNF4 α were lower

when compared to fetal (~ 49.3 RPKM) and adult hepatocytes (~68.8 RPKM), ~ 40.3 RPKM and ~ 59.8 RPKM lower; respectively (average values were used). Once again, no significant difference in variant expression was noticed between fetal and adult samples (Figure 4.9). Contrary to the pattern observed in murine hepatocytes, no switch seems to occur within fetal and adult hepatocytes in the human. This result could be due to variations in the signalling pathways between different species. There is no published data regarding the variants expression pattern within the human liver, as such, this data may provide the first real insight into the mechanisms behind HNF4 α signalling pathway in the developing human liver.

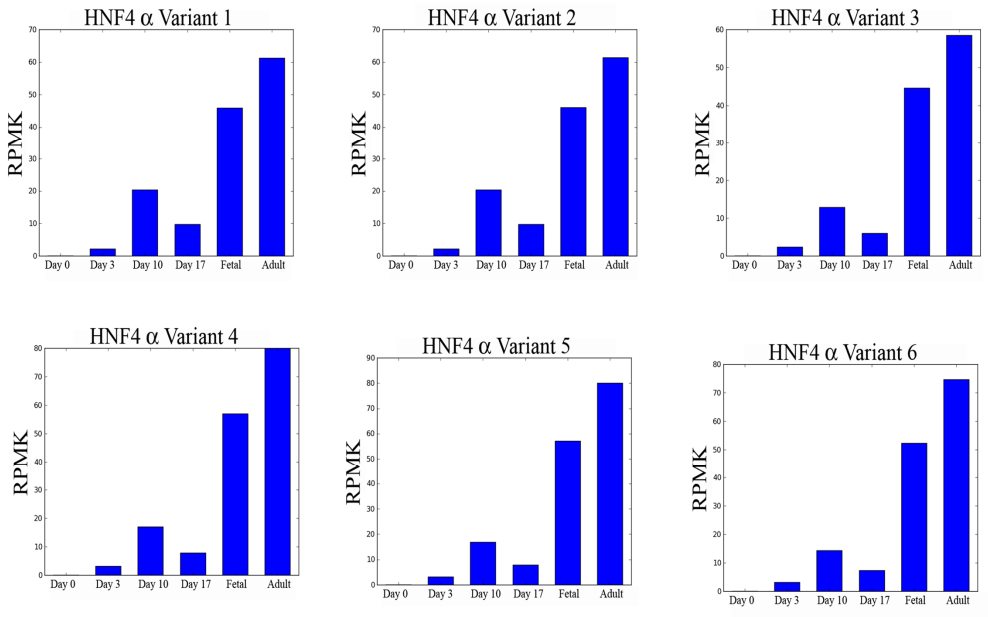


Figure 4.9 – Investigation of the Expression Profiles between the Various HNF4 α Variants using SOLEXA.
The diagram shows the expression of the HNF4 α splice variants at different time points throughout differentiation and within fetal and adult liver. As seen at a translational level, HNF4 α levels peak at the primitive/foregut endoderm stage and decrease as the cells mature to hepatocytes. However, at day 17 the HNF4 α level in all variants are significantly lower than observed in both fetal and adult hepatocytes. The differences in expression between the splice variants is approximately 20 RPMK higher in variants 4, 5 and 6 when compared to variants 1, 2 and 3 in the adult hepatocytes. A similar pattern is observed in the fetal liver, however; at day 17 HE an average of 10 RPMK is observed between all the variants.

In conclusion, the transcriptional patterning of HNF4 α matches that observed on a protein level, whereby expression levels are only comparable to fetal hepatocytes. No significant difference in expression could be detected between the HNF4 α variants in both the fetal and adult hepatocyte samples and only a slight change in expression of variant three was noted at day 17. The 'switch' between transcripts does not seem to occur in humans as demonstrated in mice, indicative of splice variant redundancy within the HNF4 α signalling cascade. Understanding the transcriptional networking of HNF4 α in hepatocytes may provide insight to its mode of regulation and subsequent function. This data suggests that HNF4 α is regulated on a translational level rather than modifying gene expression.

4.2.4 Investigating Possible Methods of HNF4 α Regulation in HE

Section 4.1.4 discusses the importance of phosphorylation on HNF4 α function and regulation. In order to manipulate HNF4 α within our system, it was important to initially understand the mechanisms behind its regulation and possible modes of function. The literature has regarded SUMOylation as a novel post translational modification altering proteins involved in a large number of cellular processes, hence, SUMOylation may play a role in HE differentiation, and more specifically in HNF4 α regulation or function.

SUMOylation is a type of post translational modification that alters the function of hundreds of proteins involved in processes such as; gene transcription, degradation, cell cycle regulation and cellular localisation. Three homologues exist in mammals, SUMO -1, -2 and -3. SUMO -2 and -3 share 95% homology with each other, but

only share 50% identity with SUMO-1 (Johnson, 2004). SUMO has been found to bind to the lysine residue on the following consensus sequence; ψKxE (where ψ corresponds to a large hydrophobic amino acid, K is a lysine residue, x is any amino acid and E is a glutamic acid residue) on the target protein (Rodriguez et al., 2001). SUMO -2 and -3 have the ability to form poly SUMO chains, covalently binding to themselves via the lysine residue at the N terminus consensus motif ψKxE. SUMO-1 lacks this consensus site and as a consequence is unable to form poly chains (Kroetz, 2005) and acts as a poly SUMO chain terminator (Ulrich, 2009). The SUMO pathway is analogous to the ubiquitin pathway in many aspects (Desterro et al., 1997). Both pathways rely on an E1 (activating enzyme), an E2 (conjugating enzyme) (Tatham et al., 2003, Desterro et al., 1999) and an E3 (the ligase), in addition to being reversible (Ulrich, 2009, Muller et al., 2001). However, there is no defined role for SUMO modification; the resulting effect is dependant on the protein modified. (Hannoun et al., 2010, Johnson, 2004, Kroetz, 2005, Muller et al., 2001)

In order to ascertain the role of SUMOylation within our system; hESCs were differentiated using standardised protocols (Hay et al., 2008) and analysed using antibodies to the SUMO, Ubc9 and SENP proteins involved in the SUMO pathway and once again, β-Actin was used as a loading control (Figure 4.10). This was carried out to validate our approach in two different media; CM (Figure 4.10 A) and MT (Figure 4.10 B).

The results from medium formulations demonstrate that hESCs in the self-renewal state have high levels of SUMO modified proteins in CM, which decrease as cells

begin to differentiate into primitive streak (Day 1 and 2), SUMO 2 modified proteins showed a similar pattern. The decrease is followed by an increase of SUMO modified proteins as the cells commit to definitive endoderm (Day 3-5), prior to HNF4 α expression.

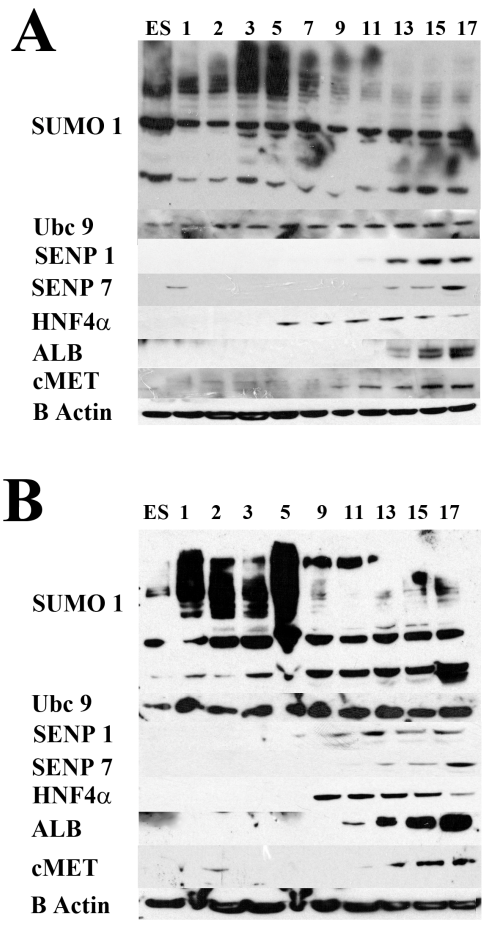


Figure 4.10 – Investigating SUMO Patterning in HE derived from hESCs Cultured in CM and MT.

Cell extracts were collected at various time points throughout the differentiation protocol and probed with SUMO1, SENP 1 and 7, HNF4 α, albumin and C Met antibodies using standard western blotting. Panel A displays HE derived from hESCs cultured in CM and panel B from hESCs cultured in MT. As previously observed, there is a decrease in SUMO modified proteins as cells adopt a more hepatic fate. This is in conjunction with the increase in the SUMO protease and mature factors such as albumin and C Met. A decrease in levels of HNF4 α is also observed. Similar patterns are observed in both medias. B-Actin was used as a loading control.

However, the initial SUMOylation pattern differs in MT cultured hESCs. The self-renewal state displays low levels of SUMO 1 modified proteins followed by a

subsequent increase through days 1 to 5, hepatic commitment, a similar pattern was noted in SUMO 2 modified proteins. The difference in SUMOylation patterns in hESCs could be attributed to the surrounding environment (Frost et al., 2011).

There is a substantial decrease in the level of SUMOylation subsequent to hepatic specification, the cells undergo maturation (days 7-17) where a dramatic decrease in SUMO modified proteins is illustrated, consistent in both media conditions. As the levels of SUMO modified proteins decrease, the levels of native SUMO increase (Figure 4.10). The increased expression of the SUMO deconjugating enzyme SENP 7 further supported this observation. Taken together, this data confirms the removal of SUMO from target proteins is achieved via SENPs activity. Peak levels of HNF4 α coincide with decreased levels of SUMOylation within the cell (Figure 4.10). In conjunction, increased levels of albumin and C-Met, indicative of hepatic maturation, corresponded to the decrease in HNF4 α, a pattern noted throughout development. In conclusion, SUMOylation seems to play a key role throughout hepatic development. Due to coinciding time points SUMOylation may also be involved in regulating HNF4 α.

Transcriptional analysis using SOLEXA was employed to understand expression levels of SUMO pathway regulators during HE differentiation (Figure 4.11). The target genes include; SUMO 1 and 2, small ubiquitin like modifiers; SAE 1 and 2, the activating enzymes (E1); Ubc 9, the conjugating enzyme (E2); PIAS 1, the SUMO ligase (E3) and three SUMO deconjugating enzymes; SENP 1, 6 and 7. Coinciding with the data generated from western blotting, a decrease in the

expression of SUMO 1/2 and SAE1/2 as well as a steady expression of Ubc9 was observed. However, unlike the protein patterning, a peak in the deconjugating enzymes, SENP 1, 6 and 7 was noted at day 3. This was followed by a decrease at day 10 and an increase at day 17, which was detected at the protein level. This could suggest that SUMO protease expression is regulated at a translational level, whereby another signal is required to activate protein translation post gene expression. In accordance with the previous data, SUMO gene expression is more comparable to fetal than adult hepatocytes (Figure 4.11).

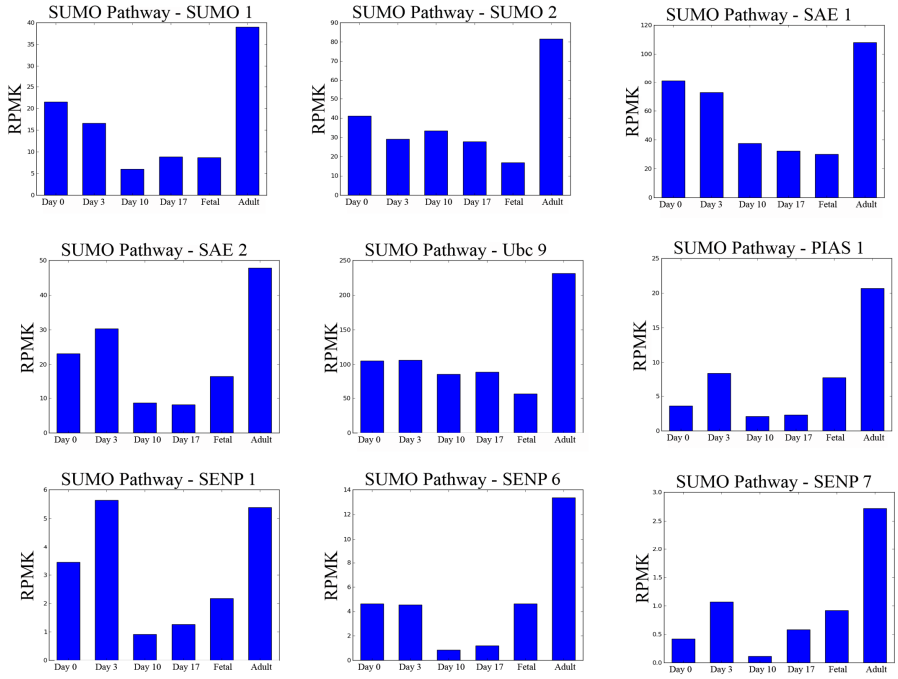


Figure 4.11 – Transcriptional Analysis of the SUMO Pathway.
In conjunction with analysing pathways at a translational level, SOLEXA was carried out to observe changes in gene expression at specific time points. The above graphs depict gene expression changes from day 0, 3, 10 and 17 hESC derived HE as well as fetal and adult human hepatocytes. The figure shows gene expression for the various components involved in the SUMO pathway. These include SUMO 1 and 2, the SUMO activating enzymes SAE1 and 2, the conjugating enzyme Ubc9, the E3 ligase PIAS I and deconjugating enzymes SENP 1, 6 and 7 (Hannoun et al., 2010, Johnson, 2004, Kroetz, 2005, Melchior et al., 2003). Once again, the gene expression patterns define day 17 HE as more fetal like than adult.

In conclusion, SUMOylation may play a role in hepatic endoderm differentiation from hESCs. The peak levels of SUMOylation at the stage of hepatic specification followed by peak levels of HNF4 α suggests a possible regulatory role of SUMOylation on HNF4 α activity.

The next chapter focuses on establishing connections between these two factors and illustrating their potential roles in HE differentiation, function and viability.

4.3 DISCUSSION

Despite the various sources of HE, they all suffer from disadvantages ranging from varied function to limited viability, which is defined by death and de-differentiation of the hepatocytes. In order to overcome these issues, efforts were we focused our efforts on understanding the biology of HNF4 α, a key transcription factor in hepatocyte biology (Odom et al., 2004). HNF4 α transcriptional activity and stability could therefore be manipulated in an attempt to fine-tune the functional and physiological nature of our differentiation model.

HNF4 α is a critical transcription factor required for normal development of the liver and pancreas. HNF4 α has been shown to regulate pancreatic β- cell proliferation, maturation of mucin producing goblet cells and crypt formations throughout development (Gupta et al., 2007). Initial studies demonstrated the loss of HNF4 α within a de-differentiated hepatoma cell line, whereby constitutive HNF4 α expression rescued albumin, fibrinogen and HNF1 protein expression, important factors produced in hepatocytes (Kuo et al., 1992). In addition, studies have shown

that HNF4 α expression occurs at 4.5 days during embryogenesis in mice, and subsequent disruption of the gene leads to death by E10.5 with significant issues noted during gastrulation. Increased cell death was indicated in the HNF4 α null embryos, which resulted in embryo lethality (Chen et al., 1994). HNF4 α has also been implicated in tumourogenesis, where the loss of HNF4 α expression is indicative of tumour progression (Lazarevich et al., 2010). HNF4 α is a liver enriched transcription factor with a large number of target genes ranging from lipid metabolism to drug detoxification. The main HNF4 α targets within the visceral endoderm include transferrin; a key regulator of iron within the body, apolipoproteins; required for lipid metabolism and Cyp 3a; involved in detoxification. Other direct targets include transthyretin; thyroxin carrier vital for maintaining homeostasis and pregnane X receptor; essential receptor for recognition of foreign toxic substances resulting in the activation of genes required for detoxification (Watt et al., 2003). Overall, HNF4 α expression is required for the efficient execution of developmental processes in addition to maintenance of metabolic functions in the adult liver.

SUMOylation patterns of hESCs cultured in CM and MT were also substantially different, whereby hESCs maintained in MT had reduced levels of SUMO modified proteins when compared to CM, figure 4.10. This data indicates the importance of standardising the hESC culture environment to generate a reliable and accurate model for further investigation. However, in both cases high levels of SUMO modified proteins were observed between days 3 and 5. This suggests the significance of SUMOylation throughout definitive endoderm formation. Further

investigation of the individual SUMOylated proteins is required to fully understand the impact of the PTM within the differentiation process. It is important to note that HNF4 α levels are required to decrease to induce maturation, however, a critical threshold value seems to be key in regulating optimal levels of maturation in addition to maintaining sufficient hepatic viability.

The levels of HNF4 α gradually decrease to a point of limited detection at day 17 on both transcriptional and translational levels (Figure 4.4 and 4.9), which may suggest an important role in HE viability. However, the decrease in HNF4 α transcription does not account for the significant loss of the protein at day 17. The decrease in HNF4 α is also observed in primary hepatocytes cultured in vitro for longer than 3 days (Tilles et al., 2002, Wege et al., 2003). This data suggests that HNF4 α is being regulated on both a transcriptional and translational level. Due to the changes in gene and protein expression, it can be assumed that translational control seems to be the more integrated form of regulation in this case.

HNF4 α has 6 known variants, each with unique time and spatial expression patterns during development and in the developed body, however, there is no published data regarding expression patterns between the fetal and adult liver in humans. As such, the expression patterns of the variants within our differentiation system and in fetal and adult hepatocytes were investigated. This was initially carried out using reverse transcription PCR; however, as this method was qualitative and not quantitative we employed SOLEXA (Figure 4.9), a highly sophisticated method of quantifying absolute numbers of gene expression. RNA was collected from various days

throughout the differentiation and the samples were processed using SOLEXA. The results depict no significant difference in expression between the HNF4 α variants in the samples or in the fetal and adult hepatocyte controls (Figure 4.9). In addition, the levels of HNF4 α expression were significantly lower in day 17 HE when compared to both fetal and adult hepatocytes. This is indicative of the fetal nature of the HE, whereby enhanced HNF4 α expression could result in improved function and viability, comparable to adult hepatocytes. Previously published data has shown a substantial difference in expression of the variants between the fetal and adult liver within the mouse. However, no differences in splice variant expression were observed between the fetal and adult hepatocytes. In turn, this data provides the first transcriptional comparison of HNF4 α splice variant expression within human development. The lack of a 'switch' is indicative of functional redundancy within the HNF4 α signalling pathway, whereby individual transcripts compensate for the activity of another. Redundancy is an inevitable consequence of evolution and in certain instances, variants are known to be redundant in developmental pathways; however, their function begins to differ depending on various metabolic processes (Pickett et al., 1995). For example, all six HNF4 α splice variants are expressed throughout human development due to the vital function of the transcription factor in gastrulation and visceral endoderm generation. However, post-development the expression of the variants is maintained but the individual transcripts regulate different cellular processes.

Several investigations have identified the various methods of regulating HNF4 α function (Section 4.14). An example includes phosphorylation, which directly alters

the transcriptional activity of HNF4 α by enhancing co- activator (Sladek et al., 1999) and DNA binding (Sun et al., 2007). Preliminary data has suggested a possible role of SUMOylation in our differentiation system, with a more specific prospect of directly modifying HNF4 α. As described earlier (Section 4.2.4), protein samples from various time points were collected and western blotted using antibodies from proteins involved in the SUMO modification pathway. Results conclude a definite pattern of SUMOylation within our differentiation model. hESCs maintained in CM displayed high levels of SUMO modified proteins, which begun to decrease as the cells differentiated towards the primitive streak. Subsequently, the levels of SUMOylation increased as the cells initiated hepatic specification. This pattern differs in hESCs cultured in MT, whereby the self-renewal state of the cells maintained low levels of SUMOylation followed by an increase throughout hepatic commitment. Differences in HNF4 α expression were also noted. These discrepancies could be a result of epigenetic changes within the cells due to differences in the culture environment. Further analysis of each media is required to fully understand their subsequent effects on hESC epigenetics. On the other hand, as the cells began to mature a similar pattern was observed in both media. As the cells adopted a more hepatic phenotype the level of SUMO modification decreased in conjunction with the decrease of HNF4 α to barely detectable levels at day 17 (Figure 4.10). This data is further supported when analysed at a transcriptional level using SOLEXA (Figure 4.11). In conclusion, SUMOylation may play a direct role in modifying HNF4 α function or regulating its stability. Subsequently, SUMOylation of HNF4 α may coordinate a mechanism essential for maintaining HE viability.

Moreover, significant differences between the transcriptional and translational levels of the SUMO deconjugating enzyme, SENP 7 were noted throughout hepatic differentiation. Gene expression of SENP 7 increased as the hESCs began to differentiate, followed by a decrease in expression at day 10, initiation of hepatic maturation, with a subsequent increase as the hepatocytes mature (Figure 4.11). In contrast, SENP 7 protein was only detected at day 11 in MT and 13 in CM with a gradual increase as the cells developed a more hepatic phenotype (Figure 4.10). The differences between transcription and translation can be attributed to post mRNA processing targeted degradation, whereby the day 3 SENP transcripts are not translated due to a number of inhibitor mechanisms. These include RNAi mediated mRNA degradation (Sonenberg et al., 2009), repressor proteins binding the 5' capped sequence and the 3' UTR on mRNA preventing ribosomal binding (Sonenberg et al., 2009), interfering with the 5' cap recognition and reducing mRNA stability by cleavage of the 3' UTR resulting in de-adenylation and subsequent degradation (Decker et al., 1994). Gene expression was re-initiated at day 10 resulting in translation detected by SOLEXA with a gradual increase in both transcriptional and translational expression levels as the hepatocytes mature. This observation coincided with the decrease in SUMO modified proteins indicative of active de-SUMOylation of target proteins throughout hepatic differentiation.

In conclusion, creating a physiological environment for hESC differentiation will allow us to unravel signalling pathways involved in hepatic differentiation. HNF4 α was identified as a potentially strong candidate to enhance HE viability required to support long-term hepatocyte culture. Previous investigations have defined a number

of regulatory mechanisms that affect HNF4 α stability and function. These findings suggest a potential role for SUMOylation in HNF4 α regulation. The next chapter focuses on defining the role of HNF4 α SUMOylation within our differentiation model. Furthermore, a mechanism may be elucidated, which in turn could be utilised to enhance HE stability and survival in culture. Subsequently, this tool could be translated for use in other hepatic cell types such including adult hepatocytes; each contributing to different clinical and research applications.

CHAPTER FIVE

DEFINING THE RELATIONSHIP BETWEEN HEPATOCYTE NUCLEAR FACTOR 4 α AND SUMOYLATION IN HEPATIC ENDODERM DIFFERENTIATION

CHAPTER 5

DEFINING THE RELATIONSHIP BETWEEN HEPATOCYTE NUCLEAR FACTOR 4 α AND SUMOYLATION IN HEPATIC ENDODERM DIFFERENTIATION

5.1 INTRODUCTION

Chapter 4 reviewed the importance of HNF4 α within our differentiation system. As a result, the potential modes of regulation were investigated and it was concluded that a possible relationship between SUMOylation and HNF4 α exists. This chapter focuses on elucidating the exact relationship between the post-translational modifier, SUMO, the vital hepatic transcription factor, HNF4 α, which may in turn prove to be a useful method of stabilizing HE viability for long-term culture, in vitro.

SUMOylation is a post-translational modification (PTM) and has been known to modify a large number of proteins involved in cellular processes such as transcription, degradation, cellular localization, cell cycle regulation and differentiation. The SUMO modification pathway is dynamic whereby protein SUMOylation is a reversible process. However, the effect of SUMOylation is dependant on the target protein. The SUMOylation pathway relies predominately upon two enzymes; the SUMO activating enzyme complex, SAE1/2 and the conjugation enzyme, Ubc9, whereby SUMO is bound via an isopeptide bond on the lysine residue in the ψKxE consensus sequence. In vivo, specific substrates require

re-localization to the nucleus before SUMO conjugation (Rodriguez et al., 2001). Studies have shown, using specific substrates, that the E3 ligase, essential in ubiquitination, is not necessary for the formation of the isopeptide bond between SUMO and the target protein (Desterro et al., 1999). However, this is not the case in all situations and is dependant on the target protein.

5.1.1 The Roles of SUMO Modification Throughout Development

SUMOylation is an essential pathway required throughout development, and its disruption results in embryo lethality (Nacerddine et al., 2005) (Nowak et al., 2006). Mouse embryos deficient in Ubc9 die at the post-implantation stage. The embryos display significant chromosomal condensation and segregation defects, which lead to the inhibition of transcription vital for normal development (Nacerddine et al., 2005). In addition, blastocysts lacking Ubc9 were viable in culture for 2 days but were not able to expand due to apoptosis within the inner cell mass. The Ubc9 mutant cells demonstrated nuclear dysmorphy and nucleoli disruption inhibiting cell cycle progression (Nacerddine et al., 2005). As a result, other models need to be developed to elucidate the mechanisms behind the SUMO pathway throughout development, such as conditional knock downs. Due to the nature of the thesis, I will specifically focus on the roles of SUMOylation in hESCs and in endodermal differentiation.

5.1.1.1 Human Embryonic Stem Cells

As mentioned in previous chapters, Oct4 is an essential transcription factor required for hESC self-renewal and pluripotency. Wei and colleagues have shown that Oct4 is SUMO modified and results in increased protein stability, DNA binding and enhanced transcriptional activity (Wei et al., 2007). The SUMO modification site is

located on lysine 118 within a SUMO consensus sequence near the N terminal DNA binding domain (Wei et al., 2007). Sex determining region Y box 2 (SOX 2) is another transcription factor required for undifferentiated hESC self-renewal and has been shown to be involved with SUMO, however further investigation is required (Van Hoof et al., 2009). SUMOylation may also indirectly affect the expression of Nanog, another vital pluripotency transcription factor. Sirtuin 1 (SIRT1) is an NAD dependant deacetylase involved in regulating p53 activity, a tumour suppressor protein vital for guarding the integrity of the genome (Han et al., 2008). Under stress, SIRT1 prevents nuclear localization of p53 induced by the production of reactive oxygen species (Han et al., 2008). This in turn, triggers apoptosis and alleviates inhibition of Nanog expression (Han et al., 2008). SUMOylation has been shown to enhance p53 transcriptional activity (Rodriguez et al., 1999), and may therefore may play a role in regulating Nanog expression.

5.1.1.2 ENDODERM

The endoderm layer is formed during embryogenesis and is the precursor of liver, pancreas, stomach, colon, urinary bladder, the lining of the urethra, the epithelial parts of trachea, pharynx, thyroid, parathyroid, lungs and intestines (Tam et al., 2003). SUMO modification is an important factor in hepatic biology as it regulates C/EBP α, a crucial factor in hepatic differentiation (Sato et al., 2006, Pedersen et al., 2001). SUMOylation of C/EBP α prevents its association with BRG1, a core subunit in the SW1/SNF chromatin-remodelling unit, which leads to the inhibition of albumin expression (Sato et al., 2006). In conjunction to this, a decrease in levels of SUMOylation as rat hepatocytes mature has also been observed (Sato et al., 2006),

suggesting an inhibitory effect of SUMOylation in hepatocyte terminal differentiation. This data is in line with the observations presented in chapters 3 and 4.

Another area of SUMO modification is its role in regulating mitochondria. The mitochondria are an essential component of hepatocytes, the main high-energy cell type in the liver, and are required for efficient liver function, due to its role as the central energy source. Mitochondrial levels within the cell are dynamic and continuously undergo fusion and fission (Twig et al., 2008, Frazier et al., 2006). An increase in SUMO-1 expression results in an increase in mitochondrial fragmentation implemented by the stabilization of the GTPase dynamin-related protein 1 (DRP1) (Harder et al., 2004). Additional investigations regarding the above pathway have revealed the importance of SENP 5, a SUMO deconjugating enzyme, in maintaining normal mitochondrial morphology and levels of reactive oxidative species within the cell, which is executed partly by SUMO deconjugation of DRP1 (Zunino et al., 2007).

5.1.2 The Roles of SUMO Modification within the Cell

Over the last decade, a number of investigations have been carried out to establish the possible roles of SUMOylation within the cell in response to various external stimuli. SUMO has been shown to modify protein activity by regulating other PTMs such as ubiquitination and phosphorylation. For instance, the SUMOylation of IκBα, a vital element in the inflammatory response, inhibits its ubiquitination thus preventing its subsequent degradation, hindering NF-κB activation and nuclear

translocation (Desterro et al., 1998). Inactive NF-κB is maintained within the cytoplasm of unstimulated cells by IκB inhibitor proteins (Desterro et al., 1999). Once induced, IκB is targeted for ubiquitination and subsequent degradation via the 26S proteasome releasing NF-κB, allowing its translocation into the nucleus and the activation of specific genes required for the inflammation response (Hay et al., 1999). SUMOylation of IκBα prevents is degradation and retains NF-κB in an inactive state (Desterro et al., 1998). Interestingly, phosphorylation provides another level of regulation, whereby phosphorylation of IκBα is required for its ubiquitination whilst it inhibits SUMOylation (Desterro et al., 1999). Therefore, the phosphorylation state of IκBα determines its subsequent fate.

SUMOylation has also been implicated in regulating the activity of 'the guardian of the genome', the tumour suppressor protein p53 (Rodriguez et al., 1999). In healthy cells, p53 is continuously ubiquitinated and degraded (Lu et al., 1993). However, upon stress induction, levels of p53 increase within the cells resulting in the expression of specific genes vital for cell cycle arrest and apoptosis (Waldman et al., 1995). Rodriguez and colleagues have demonstrated the SUMOylated nature of p53 in response to UV exposure. In addition, SUMOylation enhances the transcriptional activity of p53, an essential characteristic for protecting the integrity of the genome (Rodriguez et al., 1999).

In response to heat shock, erythroleukemia cells induce transcription of heat shock factor 1 (HSF1). After its translation, HSF1 is phosphorylated prior to its SUMOylation, which enhances its DNA binding (Hong et al., 2001). SUMO can

also regulate protein activity by modulating its interactions with other macromolecules or proteins. Various models have been proposed such as the addition of SUMO by altering protein configuration, creating a new interaction motif affecting its function (Johnson, 2004). An interesting example of interaction motifs is arsenic induced RNF4 mediated degradation of promyelocytic leukaemia (PML) bodies. In the presence of arsenic, PML is polySUMOylated, and following the recruitment of RNF4, an E3 Ub ligase, PML is ubiquitinated and degraded (Tatham et al., 2008).

In conclusion, SUMO modification alters protein function in a variety of ways, and for this reason, this chapter focuses on the relationship between HNF4 α and SUMO.

5.2 RESULTS

The first stage of defining the relationship between SUMO and HNF4 α was to investigate whether HNF4 α can be SUMO modified in vitro. This was followed by mapping the SUMOylation site on HNF4 α and determining whether mono or poly-SUMOylation played an important role in HNF4 α biology. In vitro work was carried out as a proof of principle, which in turn was reinforced with in vivo data.

5.2.1 IN VITRO SUMO MODIFICATION OF HNF4 α

The HNF4 α protein sequence was initially analysed using the software program, Abgent SUMOplot^{TM}, for possible SUMO consensus motifs; ψKxE. Two possible SUMO motifs were located on the protein; one site was present in the amino terminus of HNF4 α (**MKKE,** residues 126-129) whilst the other site was located at

the carboxyl terminus (**AKID,** residues 364-367). Both sites are conserved between all six known human HNF4 α variants and within mice and rat variants. In order to assess whether these motifs were utilized for SUMO modification, an established in vitro SUMOylation assay was performed as previously described (Tatham et al., 2008).

Primorigen Inc. kindly provided us with recombinant human wild type (WT) HNF4 α, which was industrially produced in SF-9 insect cells due to difficulties with its expression and purification. WT HNF4 α was incubated with the SUMO conjugation machinery; including SUMO-1, E1-SAE1/2 and E2- Ubc9 for 3 hours at 37°C (Section 2.7.1). The reaction was terminated at 70°C (for 10 minutes) in Nu-Page (Invitrogen, UK) LDS Sample Buffer. This assay was initially carried out using SUMO-1 as a proof of principle. The samples were then run of Nu-Page SDS PAGE protein gels (Invitrogen, UK) and analysed using western blotting. These experiments were controlled by the inclusion of a SUMO specific protease control, SENP 1, (denoted +) which cleaved SUMO modified protein, Figure 5.1. Ran GAP, a known SUMO substrate, was used as a positive control for the reaction.

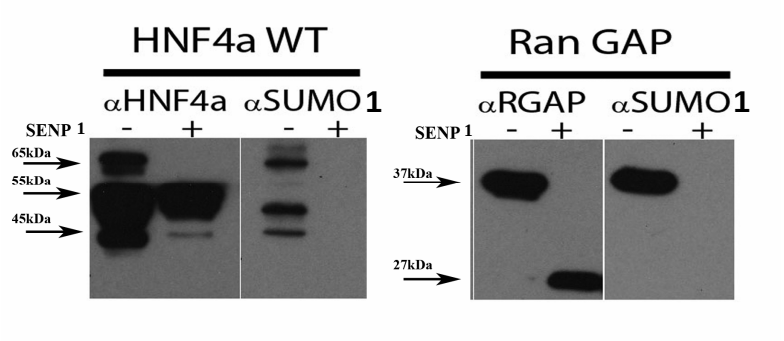


Figure 5.1 – SUMO Modification of HNF4 α, *in vitro*.
An in vitro SUMOylation assay was carried out containing the wild type form of HNF4 α in conjugation with the SUMO conjugation machinery (E1-SAE1/2 and E2-Ubc9) (Tatham et al., 2008). A control reaction which contained the SUMO deconjugating enzyme SENP1 was employed. SUMO 1 conjugation resulted in a ~10 kDa shift in protein weight, whilst SUMO deconjugation was measured by the disappearance of the band in the presence of SENP1. Ran GAP, a known SUMO substrate (Mahajan et al., 1997), was used as a positive control to ensure the efficiency of the reaction.

Figure 5.1 displays the western blots obtained from the in vitro SUMO assay. HNF4 α was successfully SUMO modified as depicted by the upward shift in the HNF4 α band (65 kDa). This shift disappeared in the presence of the deconjugating enzyme confirming SUMO modification and reaction reversibility, Figure 5.1. The blot was then stripped and re-probed with a SUMO-1 antibody to ensure the overlap between the upper HNF4 α band and the SUMO positive band, Figure 5.1. Once again, the SUMO bands disappeared in the presence of SENP 1. A similar pattern was noted in the Ran GAP in vitro assay confirming optimal reaction efficiency.

5.2.2 Mapping SUMO Modification of HNF4 α in Vitro

5.2.2.1 Expressing The N and C Terminus HNF4 α Deletion Mutants

HNF4 α possesses two potential SUMOylation sites within its structure. Deletion mutant analysis was performed to map the exact binding site. An amino (N) terminal and a carboxyl (C) terminal deletion of HNF4 α were expressed. Residues 1-139 were deleted from the N terminal mutant of HNF4 α (ΔNT HNF4 α) which resulted in the loss of the DNA binding domain, Figure 5.4. The C terminal mutant (ΔCT HNF4 α) lacked residues 350-460 which resulted in the loss of the repression domain, Figure 5.4. C and N terminal deletions of the HNF4 α were generated by amplifying the respective fragment using primer specific PCR (Section 2.1.4 for primer details). A two cycle PCR program was used where initially the specific sequence was generated to be the predominant template in the reaction (50°C for 7 cycles) followed by amplification of the sequence (60°C for 20 cycles). The PCR

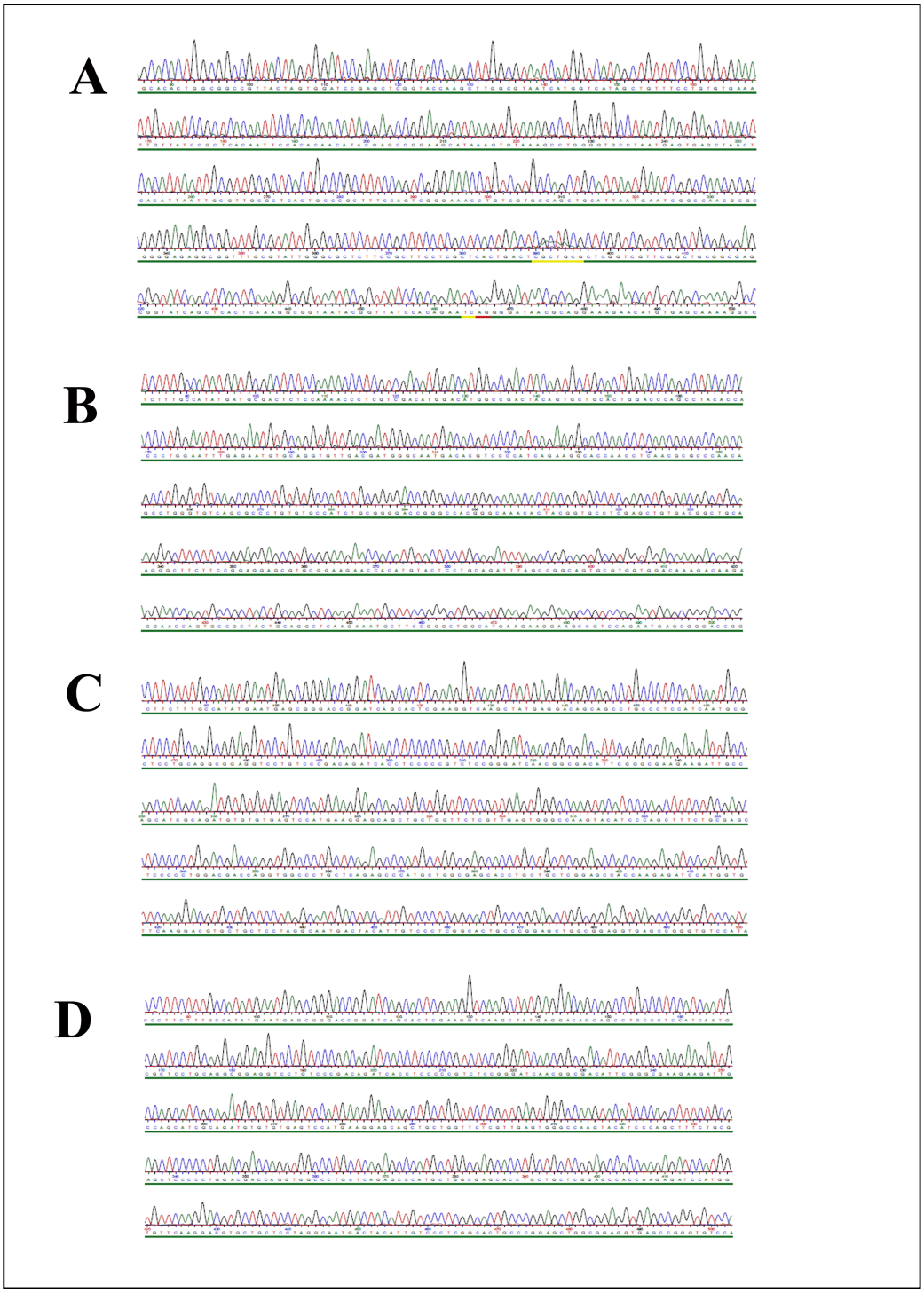


Figure 5.2 – Sequencing of HNF4 α Deletion Mutants for SUMO Mapping, *in vitro*.
These diagrams confirm the correct sequences for the deletion mutants present in the pET15b expression plasmid, providing His tags at the C terminus of the proteins. Sequencing was carried out by MWG Eurofins (Germany). **A** defines WT HNF4 α. **B.** displays ΔCT HNF4 α **C.** and **D** depict ΔNT HNF4 α and point mutant HNF4 α.

products were then run on an agarose gel to ensure successful amplification, and that the products were well amplified and ran at the expected size (~1000 bp). The DNA sequences were then gel purified and sent for sequencing, MWG Eurofins (Germany), to ensure the DNA sequence for each mutant was accurate and that no point mutations had occurred during the process, Figure 5.2.

Once the sequences of the mutants were confirmed, Figure 5.2, the fragments were cloned into the pET15b Vector (Novagen), using specific Bam HI and Nde I restriction sites. Digestion analysis was carried out to confirm the presence of the correct insert, indicated by the size of the digested bands, Figure 5.3. All mutants contained a His tag to aid in the purification of the HNF4 α deletion mutants.

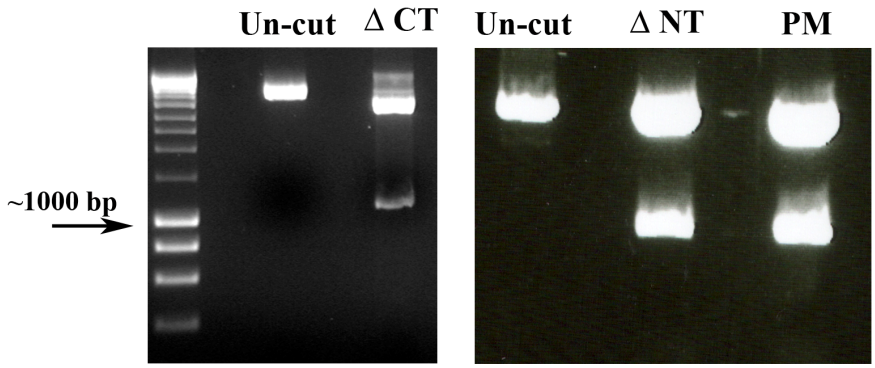


Figure 5.3 – Generation of HNF4 α Deletion Mutants for SUMO Mapping, *in vitro*.
The displayed gels show the bands corresponding to the specific deletion mutants, each at the expected base number, after digestion with BamHI and NdeI restriction enzymes, and uncut vectors were used as negative controls. The pET 15b expression plasmid was also digested and purified as the backbone vector for protein expression.

The constructs were then transformed into BL21 E-coli (Stratagene) using chemical induced transformation. Four clones were picked for each mutant and successful

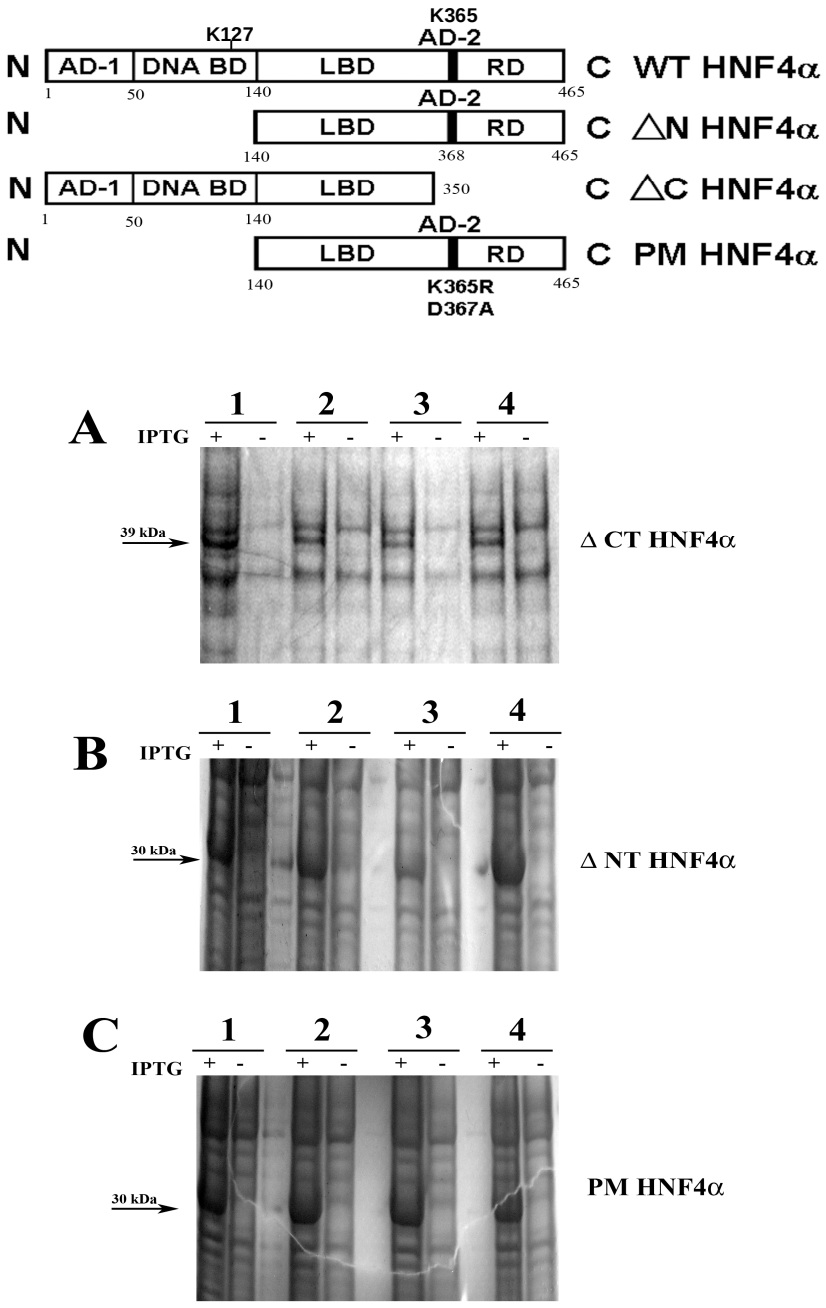


Figure 5.4 – Protein Expression of HNF4 α Deletion Mutants.

HNF4α WT and the various deletion mutants were generated in order to identify the region of SUMO modification in HNF4α. HNF4 α contains a functional domain structure which includes; the activation domains (AD-1 and AD-2), the DNA binding domain (DNA BD), the ligand binding domain (LBD) and the repression domain (RD) (Sladek et al., 1990)The mutants were induced in BL 21 cells and run on SDS gels to check for correct expression. **A.** The gels depict four different clones of BL 21 cells that contain the required plasmid with the Δ CT HNF4 α insert. Each clone was induced using IPTG, the cell pellets were lysed, run on an SDS PAGE gel and stained with Commassie Blue, in addition to a non induced control. The presence of a unique band at 39 kDa in the induced lane suggests the efficient expression of the Δ CT HNF4 α protein. This is band is not present in the non induced control. **B.** The same result occurs with the expression of the Δ NT HNF4 α protein, with a unique band present at 30 kDa as well as the point mutant (PM) HNF4 α protein (**C**).

transformation and amplification was investigated. The BL21 cells were then cultured and induced with 0.5mM IPTG (Invitrogen) to stimulate protein expression, and non-induced BL21 clones were used as negative controls. The cells were spun down and were lysed in SUMO lysis buffer. The samples of the four clones for each mutant were then run on an SDS PAGE gel and analysed for protein expression using Coomassie staining, indicated by a band present only in the group of cells that were induced, Figure 5.4.

As observed in Figure 5.4, all four clones expressed the mutant protein, which was not present in non-induced cell controls. ΔCT HNF4 α was successfully expressed at 39 kDa, Figure 5.4 A, with high efficiency in clones 1 and 4. ΔNT HNF4 α was also optimally expressed at 30 kDa, Figure 5.4 B, with highest efficiency in clone 4. As a result, clone 4 for both the ΔCT and the ΔNT HNF4 α mutants were used for the remainder of the investigation. A large-scale induction was set up (400ml culture) to allow for sufficient protein production, Figure 5.5 A. The induced cells were pelleted and lysed using SUMO lysis buffer. Samples were collected, run on an SDS PAGE gels and Coomassie stained to ensure adequate protein expression, Figure 5.5 A. As noted in Figure 5.5 A, sufficient amounts of protein were produced for both mutants. The samples were then spun down using a high-speed centrifuge to separate the soluble and insoluble fractions, and stored at -80°C for later use.

5.2.2.2 The Purification and Characterization of The N and C Terminus HNF4 α Deletion Mutants

The second stage of expressing the ΔCT and the ΔNT HNF4 α mutants was to purify the individual proteins. Samples were collected from both the soluble and insoluble

fractions and were analysed using western blotting. The samples were probed with both His tag and HNF4 α specific antibodies in order to identify which fraction the protein was expressed in. It was found that both HNF4 α mutants were expressed in the soluble fraction. As such, the soluble fraction obtained from the induced bacterial cells for each mutant was purified using nickel agarose beads that specifically pull down His tagged proteins. The soluble fractions were then incubated with the nickel-agarose beads to allow specific

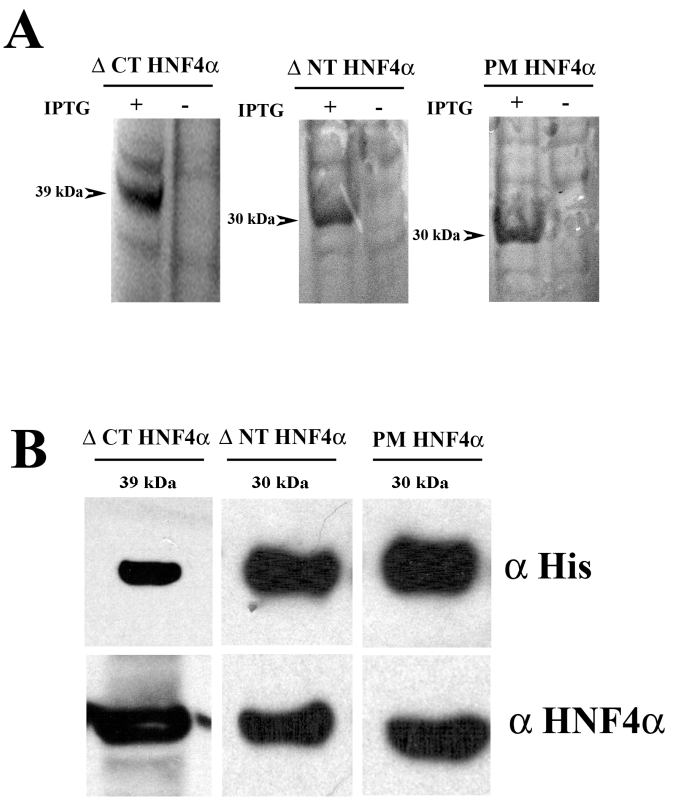


Figure 5.5 – The Large Scale Production and Purification of the HNF4 α Deletion Mutants.
A. Unique bands present in the induced lanes at 40 kDa and 35 kDa suggest the expression of the Δ CT HNF4 α, Δ NT HNF4 α and the PM HNF4 α proteins, respectively, during large scale protein production. **B.** The cell pellets were lysed and purified using nickel beads that specifically pull down His tagged proteins. The resulting eluates were run on SDS gels and probed with HNF4 α and His Tag specific antibodies. Specific bands at the correct molecular weights were observed for the Δ CT HNF4 α, Δ NT HNF4 α and the PM HNF4 α proteins. This confirms the successful expression of the required HNF4 α deletion mutants.

binding of the His tagged proteins. The columns were washed thoroughly to remove all non specific binding and improve protein purity. The His tagged HNF4 α protein mutants were then eluted and collected in 50 µl aliquots over 20 eppendorfs. A Bradford assay was used to identify the samples that contained the majority of the eluted protein. The positive samples were then pooled and dialysis was carried out to further purify the proteins. The proteins were run on an SDS PAGE gel and analysed for the HNF4 α proteins using western blotting, Figure 5.5 B.

The bands displayed in Figure 5.5 B are positive for the His tag and HNF4 α at the expected sizes, 39 kDa for the C terminal deletion and 30 kDa for the N terminal deletion. This confirms the successful expression and purification of high quality ΔCT and ΔNT HNF4 α proteins, each containing a His tag. The HNF4 α mutants were aliquoted and stored at -80°C for later use.

5.2.2.3 In Vitro SUMO Modification of The N and C Terminus HNF4 α Deletion Mutants

The C and N terminal HNF4 α deletion mutants were expressed, purified and characterized successfully. An in vitro SUMO assay was carried out on each mutant to map out the specific SUMOylation site, as previously mentioned in section 5.2.1. Ran GAP, SENP 1 and WT HNF4 α were used as additional controls for the reaction. The reactions were arrested after 3 hours and the samples were run on SDS PAGE gels and the results were analysed using western blotting.

Figure 5.6 shows efficient SUMO modification noted by the shift in WT HNF4 α (65 kDa) and Ran GAP proteins (37 kDa) (upper bands) which disappear in the presence

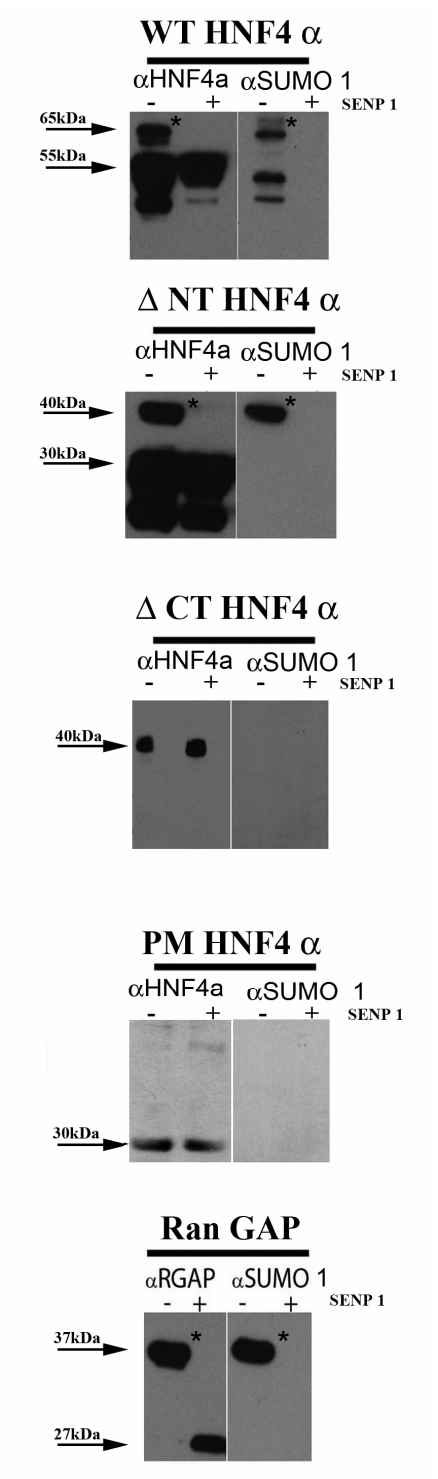


Figure 5.6 - Mapping SUMO Modification of HNF4*α*, *in vitro*.
An in vitro SUMO assay was carried out to confirm the SUMOylation site within the HNF4 α protein. The use of the deletion constructs and point mutant (Lysine 365 to Arginine 365 and Aspartic acid 367 to Alanine 367) demonstrates that the C terminus in HNF4α is absolutely required for SUMOylation in vitro, specifically at lysine 365. A control reaction which contained the SUMO deconjugating enzyme SENP1 was employed denoted by +/-. SUMO conjugation resulted in a ~10 KDa shift in protein weight, marked with an *, whilst SUMO deconjugation was measured by the disappearance of the band in the presence of SENP1.

of the deconjugating enzyme SENP 1. A similar pattern in the band shift was observed in the ΔNT HNF4 α mutant protein (40 kDa, a 10 kDa shift upwards), Figure 5.6. The upper bands in all three cases were also positive for SUMO-1. However, no observable shift was noted in the ΔCT HNF4 α mutant protein (40 kDa), Figure 5.6, and is thus not SUMO modified. This confirms the SUMO modification site is located in the C terminus of HNF4 α.

In order to confirm the absolute requirement for the carboxyl terminal consensus site in the SUMO modification process, conservative point mutagenesis (PM) of the SUMO consensus motif Ψ-K-x-D/E between residues 364-367 at the carboxyl terminus was performed. The lysine residue at position 365 was mutated to an arginine residue and the aspartic acid residue 367 was mutated to an alanine residue using single site directed mutagenesis. Two base pair mutations were generated to fully obliterate the ability for the site to be SUMO modified. This was implemented using the Quick Change II Site Directed Mutagenesis Kit (Stratagene). The mutated sequence was verified, Figure 5.2, and cloned into the pET 15b vector, which was then transformed into XL-1 Blue supercompetent cells using heat shock. Positive clones were identified using blue-white selection followed by small-scale protein expression and verification, Figure 5.4. The protein was then expressed on a large scale, Figure 5.5, purified, Figure 5.5, and employed in the *in vitro* SUMO assay, Figure 5.6, as previously described. SUMO modification (30 kDa) on the HNF4α PM was not detected, indicating that HNF4α SUMOylation proceeds through the SUMO consensus site at the carboxyl terminus on lysine residue 365, in vitro.

5.2.3 INVESTIGATING SENP PROTEIN SPECIFICITY IN SUMO-1 DECONJUGATION

There are seven known isoforms for the SUMO specific isopeptidases; including SENP 1, 6 and 7 (Melchior et al., 2003). SENPs contain a Ulp domain, located at the C terminus and are responsible for cleaving the isopeptide bond between SUMO and the target protein (Mukhopadhyay et al., 2007). Each individual SENP has a unique N terminal sequence that alters its cellular localization. This suggests that the SENPs have a distinct set of substrates under their regulation (Mukhopadhyay et al., 2007).

Next, the deconjuation specificity of SENP 1, 6 and 7 in cleaving the isopeptide bond between SUMO-1 and WT HNF4 α was investigated.

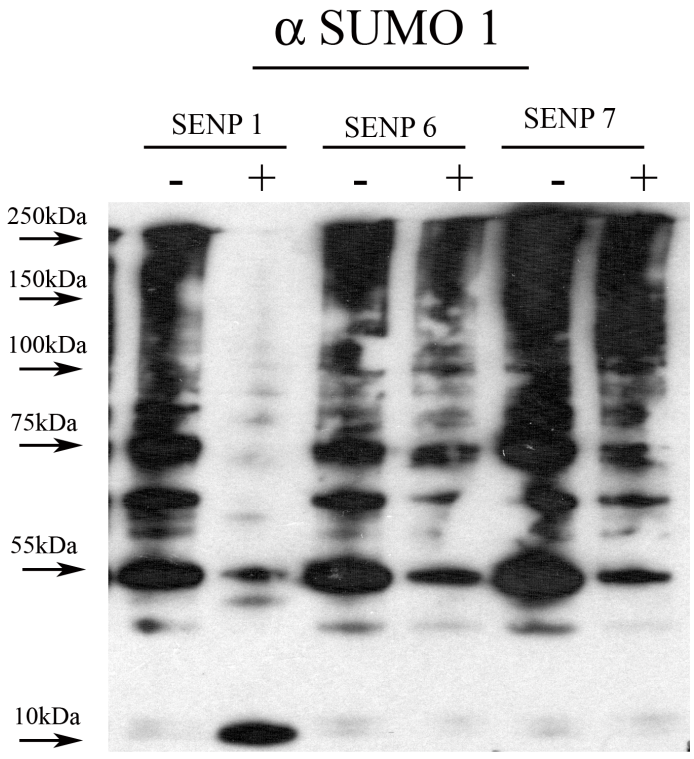


Figure 5.7 – Determining SUMO-1 Deconjugating Enzyme Specificity.
There are several known SUMO specific deconjugating enzymes known as SENPs (Mukhopadhyay et al., 2007, Hannoun et al., 2010). To establish if these SENPs have differing specificities, in vitro SUMOylation assays were carried out using SUMO 1 and subsequent deconjugation with SENPs 1, 6 and 7 were performed. The results show that SENP1 was the most efficient deconjugating enzyme to cleave the isopeptide bond between SUMO 1 and the target protein, HNF4 α. The use of SENP 1 resulted in the highest level of free SUMO (10 kDa band) when compared to SENP 6 and 7.

Professor Hay, (The University of Dundee), kindly provided the recombinant SENP proteins. An in vitro SUMO assay was carried out using WT HNF4 α as the target protein and deconjugation assays were performed using three different SENPs and the samples were analysed using western blotting. Figure 5.7 depicts the varied specificity between each SUMO specific protease. The reactions that contain the SENPs are denoted by a + sign. SENP 1 had the greatest level of specificity when deconjugating HNF4 α, with no conjugated HNF4 α remaining. SENP 6 and 7 possessed limited deconjugating capabilities, distinguished by the residual conjugated forms of HNF4 α. SENP 1 was also able to recycle the deconjugated SUMO-1 protein back into the reaction, noted by a positive band for SUMO-1 at 10 kDa. The SUMO-1 proteins deconjugated by SENP 6 and 7 were either still bound to the SENP protein itself or to another component in the reaction, and therefore, no free SUMO-1 was noted in the reactions. Overall, SENP1 is the most efficient isopeptidase for deconjugating SUMO-1 from HNF4 α. (Kaikkonen et al., 2009, Lima et al., 2008, Dou et al., 2011)

5.2.4 Investigating SUMO Modification of HNF4 α in Vivo

HNF4 α is SUMO modified in vitro at its C terminus specifically on lysine residue 365. Consequently, the SUMOylation state of HNF4 α within our differentiation system was examined. hESCs were differentiated using our standardized protocol and protein samples were collected at various time points. The protein extracts were subsequently used for immuno-precipitation to specifically pull down the endogenous HNF4 α protein. This was achieved by cross-linking the HNF4 α antibody to protein A beads (Sigma Aldrich, UK), incubating the extracts with the

cross-linked beads and eluting the bound proteins. The eluate was then analysed by western blotting.

5.2.4.1 CROSS LINKING THE ANTIBODIES

The HNF4 α antibody was cross-linked to protein A sepharose beads using dimethyl pimelimidate and ethanolamine. Cross linked and non-cross linked antibodies and beads were incubated with Nu PAGE LDS sample buffer at 70°C for 10 minutes to

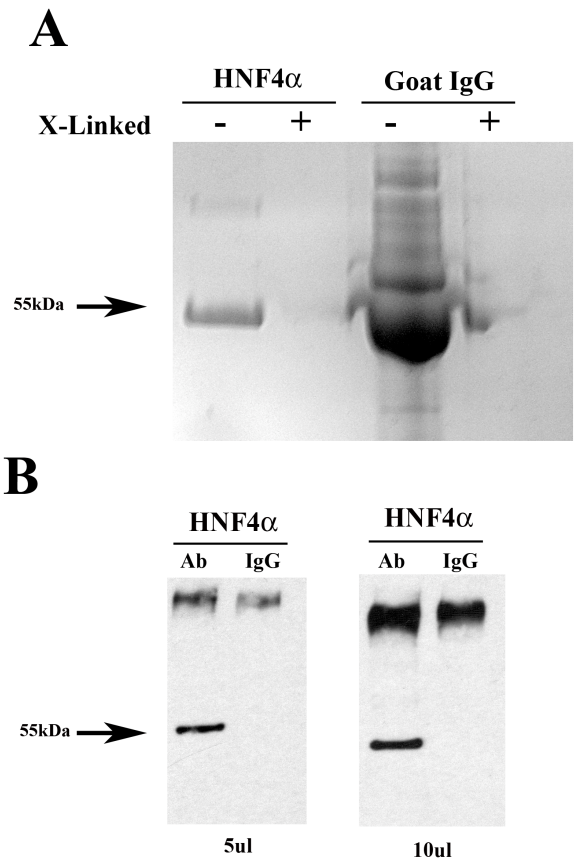


Figure 5.8 – Production of Sepharose Beads Cross Linked with the HNF4 α Antibody.
A. The disappearance of the 55 kDa band in the cross linked lane suggests the successful cross linking of the HNF4 α and IgG antibodies to the Protein A sepharose beads. **B.** The beads crossed linked to the HNF4 α antibody were used in a test pull down experiment using cell extract from the hepatocarcinoma cells line, C3A's. As observed in the blots, a band at 55 kDa is positive for the HNF4 α antibody, which is not present when using non-cross linked beads. The band intensifies as the amount of beads used in the pull down increases to 10ul. This data confirms the specificity and efficiency of the HNF4 α cross-linked protein A sepharose beads.

break all non-covalent bonds and linearise the proteins. The samples were run on SDS PAGE gels and Coomassie stained. This technique provides evidence for cross-linking efficiency. As observed in Figure 5.8 A, no bands were present in the cross linked samples suggesting that the antibody has been covalently attached to the sepharose beads. The non-cross linked samples had a band present at approximately 60 kDa, indicating the presence of the HNF4 α antibody, Figure 5.8 A. A goat IgG antibody was also cross-linked and examined, representing a negative control for the immunoprecipitation (IP).

The Coomassie stained gel shows the successful cross-linking of the HNF4 α and IgG antibody to the protein A sepharose beads. The next step was to carry out a test IP to investigate the efficiency of the cross-linked beads. C3A cells, a sub clone of the hepatoblastoma HepG2 cell line, were lysed in SUMO lysis buffer and were incubated with the HNF4 α and IgG cross-linked sepharose beads. The beads were washed thoroughly and the bound proteins were eluted. The eluates were then run on an SDS PAGE gels and analysed using western blotting.

Figure 5.8 B substantiates the ability of the cross-linked sepharose beads to specifically pull down the HNF4 α protein from cell lysate. A no antibody control was used to account for background staining and non-specific binding. The volume of sample loaded was doubled to demonstrate the linear increase of the HNF4 α protein detected, which was found to be a 1.5 fold increase (carried out by densitometry analysis). In conclusion, the HNF4 α and IgG antibodies were cross

linked effectively to protein A sepharose and were capable of specifically precipitating HNF4 α from cell extracts.

5.2.4.2 IMMUNOPRECIPITATION OF HNF4 α

Immunoprecipitation was employed to identify the SUMOylation state of the HNF4 α protein within the differentiation model.

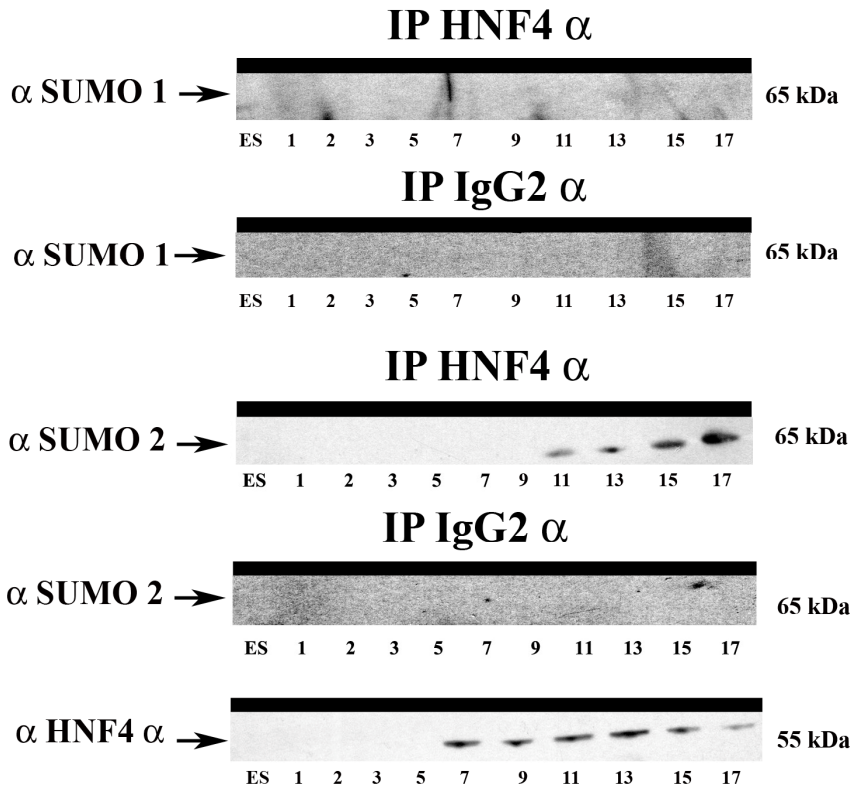


Figure 5.9 - SUMO Modification of HNF4 α, *In Vivo*.
hESCs were differentiated to hepatic endoderm and samples were prepared at the days indicated (Hay et al., 2008). HNF4 α was pulled down using HNF4 α and an IgG control antibody covalently cross linked to Protein G sepharose beads. Following incubation and extensive washing; the beads were eluted, separated by SDS PAGE and Western blotted and probed for SUMO 2. An increase in HNF4 α protein modification by SUMO-2 was observed as differentiation progressed. The negative control IgG demonstrated that the assay was operating specifically. The increase in levels of SUMO 2 modification on HNF4 α is timed with the decrease in the levels of HNF4 α throughout hepatocyte differentiation. As such, SUMO modification of HNF4 α may affect its subsequent stability.

Protein extracts were collected at various time points throughout differentiation and were individually incubated with the antibody cross-linked beads. This resulted in the specific purification of the HNF4 α protein from these extracts, an IgG antibody was used as a negative pull down control. The immunoprecipitated complexes were separated using SDS PAGE electrophoresis, western blotted and probed for SUMO 1 and 2. Differentiating between SUMO 1 and 2 modification may provide insight into the mode of regulation and possible mechanism of function.

HNF4 α was indeed modified by SUMO 2 within the differentiating cells, Figure 5.9. No detectable levels of SUMO 1 modification were observed on the HNF4 α protein. The results were specific to HNF4 α as no bands were present in the IgG pull down control as expected, Figure 5.9. Interestingly, despite the gradual decrease of HNF4 α throughout differentiation, Figure 5.9, an increase in levels of SUMO 2 in the immunoprecipitation experiments was noted. This suggests that SUMO 2 may be involved in regulating HNF4 α protein stability in vivo. Interestingly, multiple HNF4 α bands defining poly-SUMOylation were not noted, this could be due to degradation of the poly-SUMOylated chain during the IP process. In conclusion, HNF4 α is SUMO-2 modified in vivo, most likely poly-SUMOylated, which in turn may play a role in regulating protein stability.

5.2.5 Elucidating the Effect of SUMO Modification on HNF4 α

5.2.5.1 The Involvement of RNF4 in HE Differentiation

Recent developments in the field have suggested that RNF4 and poly-SUMOylation play a role in regulating protein stability. Tatham and colleagues have proved that

poly-SUMOylation of the promyelocytic leukaemia (PML) protein in the presence of arsenic resulted in its subsequent ubiquitination and degradation via RNF4 (Tatham et al., 2008). It should be noted that PML is only poly-SUMOylated in the presence of arsenic. RNF4 is an E3 ubiquitin ligase with specialized regions known as SUMO interaction motifs (SIMs) (Sun et al., 2007). RNF4 is vital throughout development as deletion mutants result in embryonic lethality and demonstrate increased levels of DNA methylation (Hu et al., 2010). RNF4 contains a conserved RING domain responsible for its ligase properties and four SIM domains required for poly-SUMOylation recognition (Sun et al., 2007). SIM motifs consist of a hydrophobic core attached to a stretch of acidic or phosphorylated residues aligned as a β strand providing a complementary SUMO binding region (Ulrich, 2008). Mutation of the SIM regions progressively hindered poly-SUMO chain binding (Tatham et al., 2008). It was also found that RNF4 could not distinguish between SUMO 1 and SUMO 2 chains. A mutant SUMO 1 protein capable of forming SUMO chains was generated and used for poly-SUMOylation, which in turn was recognized by RNF4. Hence, RNF4 is solely capable of detecting and interacting with poly-SUMO chains with no preference to the isoform present (Tatham et al., 2008). However, SUMO 1 is unable to form chains in its native environment (Johnson, 2004).

In response to this body of work, the role of the RNF4 within our differentiation system was investigated. RNF4 expression was detected in vitro from days 5 to 17 throughout differentiation, Figure 5.10. RNF4 protein expression increased gradually throughout differentiation, which coincided with the decrease in HNF4 α expression and the increase in levels of SUMO-2 modified HNF4 α. This observation is compatible with the idea that poly-SUMOylation and RNF4 regulate HNF4 α

stability. RNF4 expression was detected in adult hepatocytes, which could suggest the involvement of RNF4 in maintaining the terminally differentiated state of mature hepatocytes and it may contribute to tissue homeostasis.

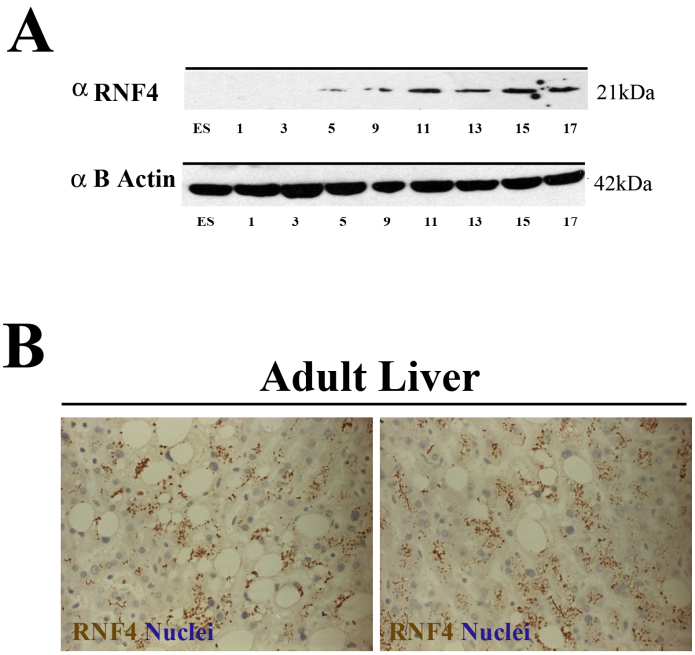


Figure 5.10 –RNF4 Patterning Within Hepatic Differentiation.
A-hESCs were differentiated to hepatic endoderm and protein samples were prepared at the days indicated (Hay et al., 2008). The samples were then run on an SDS PAGE gel and analysed for RNF4 using western blotting. A gradual increase in RNF4 expression is noted from day 5 onwards, peaking at day 17 in conjunction with peak levels of SUMO modified HNF α obtained from IP experiments. This coincides with the decrease in levels of endogenous HNF4 α. This data supports the role of RNF4 and SUMO-2 in regulating HNF4 α stability. **B-** Sections were DAB stained for RNF4 to analyse the spatial patterning within the adult liver. The sections were positive for RNF4 as indicated by brown staining. The images were taken at x40 magnification. However, due to discrepancies between various sections no conclusions could be drawn.

In conjunction with the above experiments; adult liver sections were also DAB stained using enzymatic antigen retrieval to observe the spatial patterning of RNF4 within the adult liver, Figure 5.10 B. The sections stained positive for RNF4, indicated by the brown stained regions, Figure 5.10 B, which confirmed the western blot data noted in Figure 5.10 A.

5.2.5.2 POLY-SUMOYLATION OF HNF4 α IN VITRO

The next stage in the investigation was to assess the ability of HNF4 α to be poly-SUMOylated in vitro. The in vitro SUMO assay was employed using SUMO-2 in replacement of SUMO-1. SENP 1 was once again used as a control to ensure efficient SUMOylation. Ran GAP was also SUMOylated and used as a positive control to measure the efficiency of the reaction. Once the reactions were stopped, the proteins were run on an SDS PAGE gel and western blotted. Figure 5.11 confirms the ability of HNF4 α to be poly-SUMOylated. Multiple bands were present in both the HNF4 α and Ran GAP reactions indicating the various lengths of the SUMO chains attached. The bands disappeared in the presence of the deconjugating enzyme SENP 1, Figure 5.11. These blots were stripped and probed for SUMO 2, and a similar pattern was observed, data not shown.

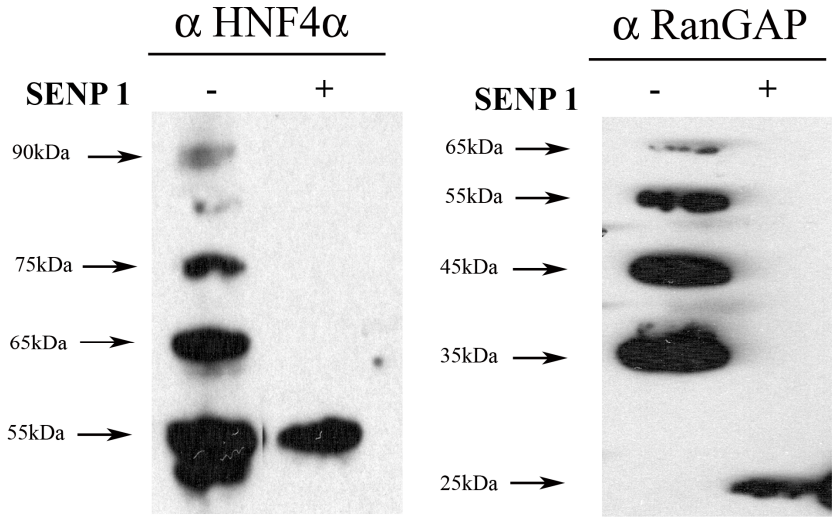


Figure 5.11 – Poly-SUMOylation of HNF4 α, *In Vitro.*
HNF4 α was poly-SUMOylated using an in vitro SUMO assay and replacing SUMO 1 with SUMO 2. The reactions were run on SDS PAGE gels and western blotted for HNF4 α. Ran GAP and SENP 1 were used as control for the reaction. Both HNF4 α and Ran GAP were successfully poly-SUMOylated as indicated by the multiple bands which disappeared in the presence of SENP 1.

5.2.5.3 INVESTIGATING SENP PROTEIN SPECIFICITY IN SUMO-2 DECONJUGATION

An in vitro SUMO assay was performed followed by deconjugation assays using SENP 1, 6 and 7 in order to measure the efficiency of each isopeptidase to cleave poly SUMO chains from HNF4 α.

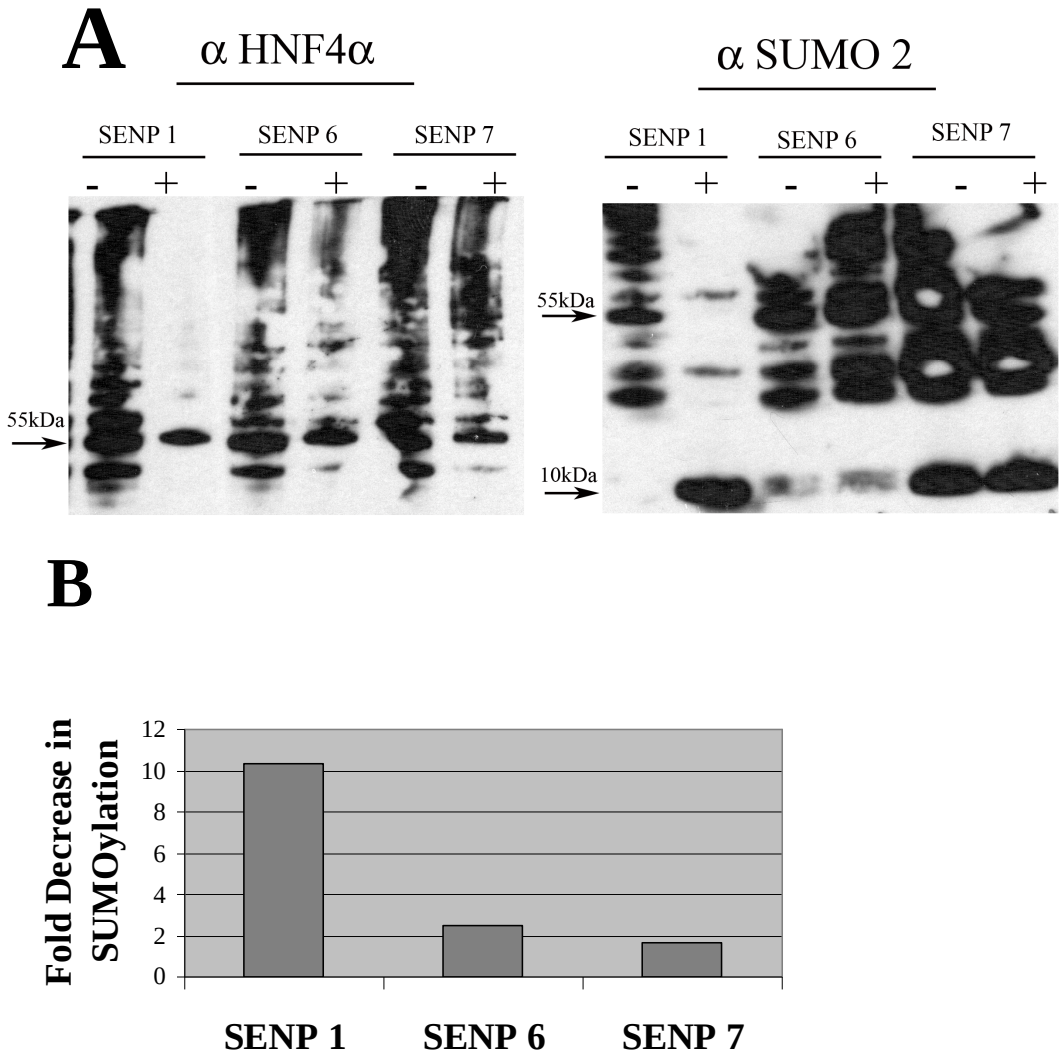


Figure 5.12 – Determining SUMO-2 Deconjugating Enzyme Specificity.
A. There are several known SUMO specific deconjugating enzymes known as SENPs (Mukhopadhyay et al., 2007, Hannoun et al., 2010). To establish if these SENPs have varied specificities, in vitro SUMOylation assays were carried out using SUMO 2 and subsequent deconjugation with SENPs 1, 6 and 7 were performed. The results show that SENP1 was the most efficient deconjugating enzyme to cleave the isopeptide bonds between SUMO 2 chains and the target protein, in this case HNF4 α. The use of SENP 1 resulted in the highest level of free SUMO (10 kDa band) when compared to SENP 6 and 7, whilst deconjugating all poly-SUMOylated HNF4 α. **B.** The graph displays the decrease in the levels of SUMO modified HNF4 α once the deconjugating enzymes were added. Densitometry analysis displays a 10 fold decrease in SUMOylated HNF4 α in the presence of SENP 1 when compared to the 2.5 fold and 1.5 fold decrease when using SENP 6 and 7, respectively.

The results were depicted in western blots, Figure 5.12. Once again, SENP 1 was able to fully deconjugate poly-SUMOylated HNF4 α whilst SENP 6 and 7 had limited cleaving capabilities. The HNF4 α blot was stripped and re-probed with SUMO 2, similar patterns were noted. Overall, SENP-1 has greatest specificity when cleaving mono and poly SUMOylated forms of HNF4 α, Figure 5.7 and 5.12 respectively. However, this may not be an accurate representation of what occurs in vivo, as the in vitro environment may promote and favour SENP 1 activity. It should be noted that the SENP 7 reaction contained free SUMO proteins in the absence of the deconjugating enzyme. This may be attributed to less than optimal SUMOylation of HNF4 α in the initial reaction (below 100% efficiency), which occurs when SUMOylating certain substrates (Johnson, 2004).

Densitometry analysis was carried out using Image J software on the lanes before and after deconjugation, Figure 5.12 B, and indicates that SENP1 is the most effective isopeptidase for deconjugating SUMO 1 and SUMO 2 modified proteins. As observed, the levels of poly-SUMOylated HNF4 α decreased by more than 10 fold in the presence of SENP 1 when compared to the 2.5 and 1.5 fold decrease when utilizing SENP 6 and 7, Figure 5.12 B. In conclusion, SENP 1 is capable of efficiently cleaving both mono and poly-SUMOylated forms of HNF4 α.

5.2.6 POLY-SUMOYLATION TARGETS HNF4 α FOR INCREASED RNF4 MEDIATED UBIQUITINATION, IN VITRO

We hypothesized that poly-SUMOylation of HNF4 α results in its ubiquitination via RNF4. In order to test our hypothesis, we set up an in vitro ubiquitination assay in the presence and absence of RNF4. Unmodified HNF4 α was compared to HNF4 α that had been SUMO-1 and SUMO-2 modified, Figure 5.13 A. The western blots confirmed sufficient mono and poly SUMOylation of HNF4 α before use in the ubiquitination assay. The substrates were then incubated with excess ubiquitin in addition to the E1 and E2 enzymes, with and without RNF4. Four hours post incubation; the reactions were immunoprecipitated to specifically isolate the HNF4 α protein. The elutes were subsequently run on SDS PAGE gels and probed for ubiquitin using western blotting.

Ubiquitination was solely detected in reactions containing RNF4, confirming the requirement of the E3 ligase for the covalent attachment of ubiquitin, Figure 5.13 B. Furthermore, our in vitro assay demonstrated that poly-SUMOylation of HNF4 α resulted in a three-fold increase in ubiquitination when compared to mono SUMOylated HNF4 α and a six-fold increase in comparison to native HNF4 α, Figure 5.13 C. To ensure specific isolation of HNF4 α, the blots were probed with RNF4, a non-covalent interaction with the target protein, Figure 5.13 B.

No bands were detected, therefore, HNF4 α was specifically pulled down with the exception of a few ubiquitin molecules due to non-specific binding of the beads. In

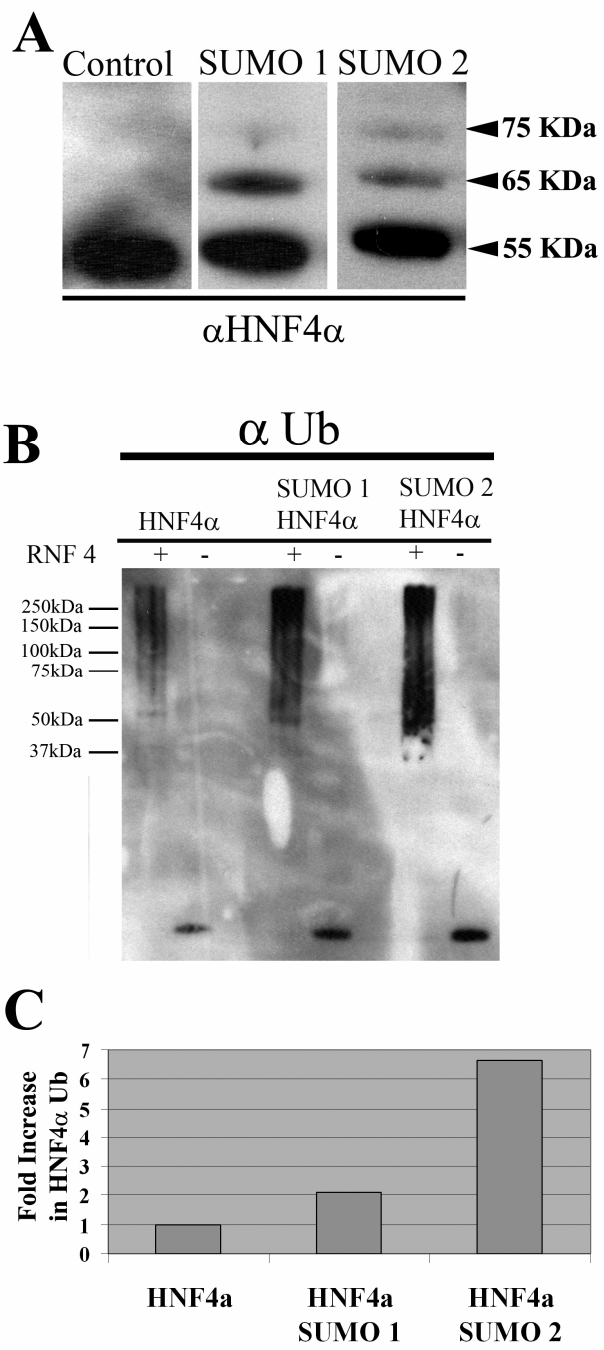


Figure 5.13 - Poly SUMOylation Marks HNF4a for RNF4 Mediated Ubiquitination, *In Vitro*.
A. In order to study the stability of SUMOylated HNF4a in vitro we employed an RNF4 Ubiqutination assay. WT HNF4α was mono and poly SUMOylated using SUMO-1 and SUMO-2 but absent from the control lane. **B.** SUMOylated HNF4α protein was incubated with ubiquitination machinery (E1 and E2) in the presence and absence of RNF4 (E3). Following assay completion, HNF4α was immunoprecipitated and ubiquitin conjugation analysed using western blotting. Ubiquitination only occurred in the presence of RNF4 and was increased when HNF4α was poly SUMOylated. No bands were positive for RNF4 suggesting specific isolation of HNF4 α. **C.** Densitometry Analysis was performed using Image J software, and the level of ubiquitination increased by 6 fold over the control when HNF4 α is poly SUMOylated when compared to 4 fold when HNF4 α is mono SUMOylated.

addition, ubiquitin positive bands for HNF4 α at approximately 40-45kDa were observed and were suggested to be HNF4 α degradation products. The degraded forms of HNF4 α were also present in previous in vitro SUMOylation assays and were positive for SUMO modification.

This data suggests that poly-SUMOylation of HNF4 α leads to preferential ubiquitination by RNF4, in vitro. Densitometry quantification was performed using the Gel Analyzer plug-in from ImageJ software. The bands were selected and equal pixels numbers were applied. A histogram was plotted using the band intensity and the area under the peak was measured accordingly. In conclusion, the above experiment has provided evidence supporting the role of SUMO modification in regulating HNF4 α stability, in vitro.

RNF4 to date is the only known ubiquitin ligase in humans that specifically recognises poly-SUMO chains via its SIM domains. The RNF4 protein is conserved throughout various species such as; *Saccharomyces cerevisiae* Slx5–Slx8 and *Dictyostelium discoideum* MIP1, whereby Slx5-8 deletion mutants result in genomic instability and a build up of SUMOylated proteins (Prudden et al, 2007). A similar phenotype is also noted in humans (Hun et al, 2010). Further investigation is required to isolate other possible ubiquitin E3 ligases containing SIM domains, which may have a function comparable to RNF4. Evidence from a number of investigations and the data presented in this chapter suggests that poly-SUMOylation may have a direct role in regulating protein degradation within cells, in conjugation with other cellular functions.

5.2.7 POLY-SUMOYLATION TARGETS HNF4 α FOR UBIQUITIN MEDIATED DEGRADATION, IN VIVO

Section 5.2.5.1 confirms RNF4 expression in our mature hepatocyte like cells, Figure 5.10 A. As such, our in vivo system contains the correct machinery required for HNF4 α modification and degradation. To gain further insight behind this process within our system, we incubated day 17 hepatic endoderm in the presence and absence of MG132, a defined proteasome 26S inhibitor. The cell extracts were collected and analyzed for HNF4 α and the SUMO proteins using western blotting.

HNF4 α was stabilized in the presence of MG132, indicating that ubiquitin dependant proteolysis is responsible for regulating levels of HNF4 α within our cells, Figure 5.14 A. Furthermore, a doublet band at ~75 kDa signifies the preservation of modified forms of HNF4 α in the absence of the 26S proteasome activity, Figure 5.14 A. We also observed a build up of SUMO-2 modified proteins in blots probed with the SUMO-2 antibody, which is compatible with our previous results, Figure 5.14 B. In addition to this, a doublet band was noted at ~75 kDa, also present in the HNF4 α blot, Figure 5.14 B. The doublet band present in Figure 5.14 A, not present in Figure 5.9, may be a result of successful preservation of poly-SUMOylation throughout the protein isolation and detection process. The IP process may have resulted in the cleavage of the poly-SUMO chain, which was avoided in this experiment as the cell extracts were immediately run on SDS page gels and probed for HNF4 α using western blotting. We then examined the stability of SUMO 1 modified proteins in the presence and absence of MG132. In contrast to SUMO 2 conjugated proteins, we observed a decrease in SUMO 1 modified proteins in

response to MG132 treatment, Figure 5.14 B. Correlating with our in vitro data; the above investigation reinforces the ubiquitin dependant regulation of HNF4 α stability within our differentiation system. The likely mechanism is as follows; an environmental cue induces poly-SUMOylation of HNF4 α resulting in the RNF4 mediated ubiquitination and subsequent degradation via the 26S proteasome. As such, SUMO modification plays a vital role in regulating HNF4 α stability in vivo.

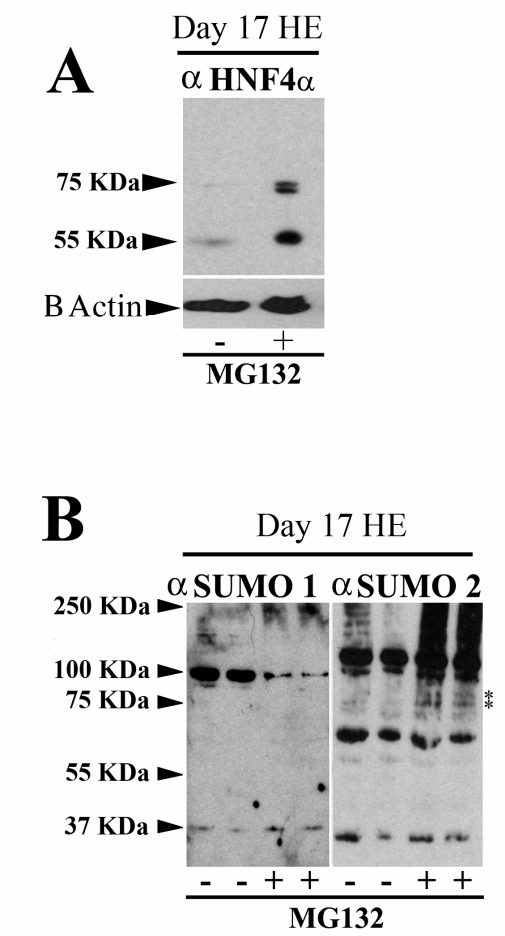


Figure 5.14 - Poly SUMOylation Marks HNF4α for Degradation, *In Vivo*.
A. In order to assess if Ubiquitin mediated proteolysis played a central role in the stability of HNF4α, we inhibited the 26S proteasome using MG132. The Day 17 hepatic endodermal cells were incubated with and without MG132 and whole cell extracts were collected, separated, Western blotted and probed for HNF4α, duplicates were run for each condition. HNF4α levels were stabilized in the presence of MG132, demonstrating the importance of ubiquitination and the 26S proteasome in the regulation of HNF4α stability. These data suggest that RNF4 mediated ubiquitination of poly-SUMOylated HNF4 α regulates protein stability in vivo. **B.** In accordance to this data, we observe a build up of SUMO 2 modified proteins in the presence of MG132, suggesting its involvement in protein degradation. The opposite is shown in SUMO 1 where there is a decrease in the level of SUMO 1 modified proteins in the presence of MG132. Overall, this data suggests that SUMO 2 may play a more global role in regulating protein degradation, which may be linked to its ability to form poly-SUMO chains. The asterisks denote bands present at 55kDa and 75kDa, similar to that observed in **A**, suggestive of poly-SUMOylated HNF4α.

5.3 DISCUSSION

In an effort to refine our differentiation system, we focused on HNF4 α, a transcription factor responsible for regulating hepatic function and viability, controlling the active expression of ~40% of genes within the liver (Odom et al., 2004). We were interested in a particular post translational modification of HNF4 α, SUMOylation, and its subsequent effect within our cells. HNF4 α regulates the formation of the visceral endoderm throughout development and is vital for gastrulation to occur (Chen et al., 1994, Watt et al., 2003). In the adult liver, HNF4 α maintains the expression of an array of genes required for efficient liver metabolism such as detoxification and tissue homeostasis (Watt et al., 2003). In addition, mutations in HNF4 α have been directly associated with pancreatic and liver diseases such as maturity onset diabetes of the young (MODY) (Ellard et al., 2006) and cancer (Lazarevich et al., 2004). Therefore, HNF4 α is an ideal candidate for refining our differentiation system.

HNF4 α has been demonstrated to be phosphorylated by protein kinase C on a conserved serine residue, which hinders DNA binding and reduces transcriptional activation in addition to enhancing its degradation (Sun et al., 2007). In conjunction, protein kinase A (Viollet et al., 1997), p38 kinase (Guo et al., 2006) and AMP activated protein kinase (Hong et al., 2003) have been demonstrated to phosphorylate the serine residue resulting in similar consequences. Acetylation, on the other hand, enhances DNA binding activity and co-activator interactions (Soutoglou et al., 2000). Hence, PTM's play a significant role in regulating HNF4 α function and stability. Initially, we found that HNF4 α was SUMO modified, in vitro, Figure 5.1. We found

2 potential SUMOylation sites; one located in the amino terminus and one found at carboxyl terminus of the protein. Through deletion and point mutation analysis we determined that the critical region for HNF4*α* modification was the consensus site found at the carboxyl terminus, Figure 5.2 - 5.6. Whilst HNF4*α* SUMOylation took place in vitro, immune precipitation experiments confirmed that HNF4α was SUMO modified in vivo, Figure 5.9. Moreover, and in contrast, to the correlation of HNF4 α and general SUMOylation pattern we observed an increase in protein SUMOylation, which resulted in decreased HNF4α stability, Figure 5.9. HNF4 α was found to by modified by SUMO-2 in vivo. SUMO-2, unlike SUMO-1, contains a consensus SUMO modification motif and is thus capable of forming SUMO chains (Johnson, 2004). SUMO-1 however, can be used as a chain terminator (Kroetz, 2005). Poly-SUMO chains were formed on HNF4 α in vitro, substantiating the in vivo data, Figure 5.11.

Literature reviews lead us to investigate the role of RNF4 throughout this process. RNF4 is a ubiquitin E3 ligase that contains specific SIM motifs responsible for recognizing poly-SUMO chains. Our results demonstrated that RNF4 expression increased as the levels of HNF4 α decreased during hepatic differentiation, Figure 5.10. As HNF4 α was capable of being poly-SUMOylated in vitro, Figure 5.11, we carried out a ubiquitination assay in the presence and absence of RNF4. Ubiquitination only occurred in the presence of RNF4, confirming its requirement in this process, Figure 5.13. Interestingly, we found that poly-SUMOylation of HNF4 α resulted in a 6-fold increase in its RNF4 mediated ubiquitination when compared to native HNF4 α and a 3-fold increase in comparison to mono-SUMOylated HNF4 α,

Figure 5.13. In order to gain further understanding of our differentiation system, were employed the use of a proteasome inhibitor, MG132. The results demonstrated that HNF4 α was stabilized in the presence of MG132, with a doublet band present at ~75 kDa, suggesting poly-SUMOylation, Figure 5.14 A. This was re-established by the increase in the levels of SUMO 2 modified proteins in samples containing MG132, with a doublet band noted at ~75 kDa, Figure 5.14 B. This suggests that SUMO-2 may in fact play a larger role in protein regulation mediated by proteasome 26S degradation. In contrast, SUMO 1 modified proteins decreased in the presence of MG132, Figure 5.14 B. These investigations support our hypothesis, Figure 5.15.

HNF4 α expression decreases throughout hepatic differentiation, which is timed with the increase in the level of SUMO-2 modification of the protein during which RNF4 expression peaks, Figure 5.10 A. We have also shown that HNF4 α can be modified in vitro by both SUMO-1 as monomers and SUMO-2 as chains, Figure 5.1 and 5.11. This data in conjunction with the literature suggests that the poly-SUMOylation of HNF4 α results in its ubiquitination and subsequent degradation mediated by RNF4. RNF4 is able to recognize poly-SUMOylation via its SIM domains, the close proximity between RNF4 and HNF4 α causes it to be ubiquitinated via the RING domain. Poly ubiquitin chains attached to HNF4 α results in its degradation via the 26S proteasome. Figure 5.15 displays the schematic of the above hypothesis. The poly-SUMOylation of HNF4 α may be caused in response to changes in cellular metabolic processes associated with differentiation.

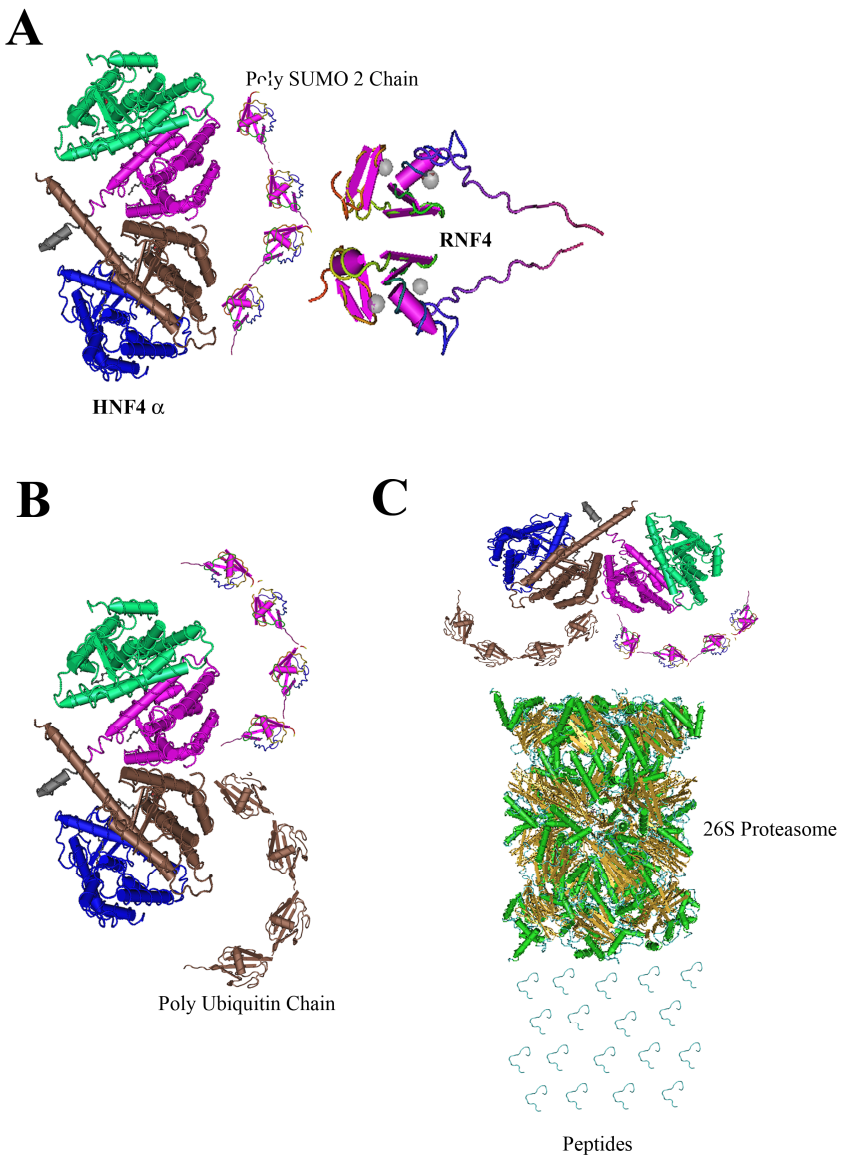


Figure 5.15 – RNF4 Mediated Degradation of poly-SUMOylated HNF4 α
This schematic represents the proposed hypothesis whereby the poly-SUMOylation of HNF4 α leads to ubiquitination and its consequential degradation mediated by RNF4. **A-** Poly-SUMOylated HNF4 α is recognized by RNF4 and binds via the SIM domains. **B-** RNF4 is an E3 ligase and binding to poly-SUMOylated HNF4 α causes its ubiquitination. **C-** Ubiquitinated HNF4 α results in its degradation via the 26S proteasome.

Studies have established similar mechanisms of regulating protein stability via poly-SUMOylation. Tatham and colleagues have demonstrated that upon arsenic

induction, promyelocytic leukaemia (PML) protein is poly-SUMOylated, which results RNF4 mediated ubiquitination followed by its subsequent degradation via the 26S proteasome (Tatham et al., 2008). Subsequent investigations demonstrated that arsenic trioxide also induced the RNF4 mediated degradation of the oncogenic fusion PML protein, PML and the retinoic acid receptor, responsible for causing acute promyelocytic leukaemia (Lallemand-Breitenbach et al., 2008). These studies provided insight into a mechanism that is translatable into a possible long-term treatment for leukaemia. Another example of RNF4 mediated protein degradation was demonstrated by Van Hagen and colleagues. Hypoxia-inducible factor 2 α (HIF-2 α), vital for regulating cell survival during hypoxia, is a direct target of SUMOylation, which results in reduced transcriptional activity of the transcription factor. Furthermore, inhibition of the 26S proteasome in conjugation with generating RNF4 knock outs indicated that poly-SUMOylation of HIF-2 α induced RNF4 mediated ubiquitination and its subsequent degradation. (van Hagen et al., 2010)

Poly-SUMOylation has also been implicated in regulating global cellular degradation, noted in Figure 5.14; whereby SUMO 2 modified proteins accumulate on proteasomal inhibition. Uzunova and colleagues has also observed a similar SUMOylation pattern in both human cells and yeast (*Saccharomyces cerevisiae*) when the proteasome has been inhibited. Additionally, when RING finger proteins where knocked down in the ubc4 and 5 yeast cell lines, levels of poly-SUMOylated proteins increased (Uzunova et al., 2007). Taken together, the results are indicative of a coordinated role between SUMOylation and ubiquitination in mediating the degradation processes within eukaryotic cells.

HNF4 α is SUMOylated at lysine residue 365, which is located in the activation domain-2 (AD-2). The AD-2 region alters its conformational state to regulate the activity of the ligand-binding domain, which adopts two different conformations depending on AD-2. As such, we have shown that the SUMOylation of HNF4 α at the AD-2 domain plays an important role in HNF4α stability and transcriptional activation during hESC stem cell differentiation. Furthermore, HNF4 α is highly dependant on co-activator interactions to enhance DNA binding and transcriptional activation (Watt et al., 2003, Hadzopoulou-Cladaras et al., 1997, Sladek et al., 1990). SUMOylation of HNF4 α may alter its ability to bind co-activators and DNA affecting its transcriptional capabilities. During hepatocyte differentiation HNF4 α was progressively modified with SUMO-2 which peaked at day 17 which, coincided with mature HNF4 α transcriptional activity, confirmed by increased serum protein production of direct HNF4 α gene targets (Figure 4.5), and reduced protein stability. However, post day 17 the HNF4α was not detected and the hepatocytes enter a program of de-differentiation.

In conclusion, HNF4 α is a crucial transcription factor required for hepatocyte function and viability. Within our differentiation system, levels of HNF4 α decrease initially which may activate the expression of hepatic genes responsible for inducing maturation. The decrease in HNF4 α is observed throughout development, however, the levels of HNF4 α are stabilized in the adult liver due to its involvement in regulating hepatic function (Cereghini, 1996, Costa et al., 2003, Lemaigre et al., 2004). As a result, the reduced viability within our differentiation system may be

attributed to decreased levels of HNF4 α, which in turn has been indicated to be regulated by poly-SUMOylation resulting in RNF4 mediated ubiquitination followed by its degradation. This mechanism has been observed in other proteins such as HIF-2 α (van Hagen et al., 2010) and PML (Tatham et al., 2008). Therefore, maintaining the balance between HNF4α stability and mature transcriptional activity is a key regulatory process in differentiation and maintenance of cell phenotype. The next chapter will focus on the roles of HNF4 α SUMOylation in hepatic endoderm differentiation using our cellular model.

CHAPTER SIX

THE EFFECT OF SUMOYLATION AND HNF4 α IN LIVING SYSTEMS

CHAPTER 6

THE EFFECT OF SUMOYLATION AND HNF4 α IN LIVING SYSTEMS

6.1 INTRODUCTION

HNF4 α is an important factor required throughout hepatic endoderm differentiation and poly-SUMOylation seems to play a role in the process. The focus of chapter 6 was to alter the expression levels of HNF4 α and observe its subsequent impact on hepatic function in living systems.

6.1.1 TECHNIQUES EMPLOYED FOR PATHWAY MANIPULATION

The two most effective methods for manipulating pathways and studying their respective consequences are pathway inhibition or stimulation using specific chemical substrates and genetic modification of the pathway using over expression and knock down systems. We carried out a thorough investigation of the field and concluded that HNF4 α had no known substrates that specifically inhibit or promote its expression and functional activity. Due to the limited knowledge and investigation of the SUMO pathway, we were unable to identify proteins or chemical molecules that explicitly targets SUMOylation. As a result, we employed the use of techniques capable of modifying the transcriptional levels of the target proteins.

6.1.2 GENETIC ENGINEERING

There are a number of successful molecular genetic approaches that result in the inhibition or stimulation of target genes. Gene targeting via homologous recombination is a popular approach as it specifically alters the native gene within the DNA without affecting other sequences. This technique can introduce genetic mutations, insert reporter genes, alter the promoter sequence or delete a gene. Plasmid vectors are another way of altering gene expression by introducing the modified DNA sequence in a host plasmid before transforming it into the target cells. This method results in transient modifications and is usually utilized in bacterial cells. However, both these techniques rely on external methods to introduce the recombinant DNA into the cells. These include electroporation; creating miniscule pores in the cell membrane to allow the DNA to enter the cell, microinjection; this method is used to introduce new genetic material without biological vectors and most recently a new method referred to as bioballistics was developed. Bioballistics involves the use of metal slivers coated with the genetic material in interest, which is then 'shot through' the cells using a shotgun. The metal sliver passes through the cell leaving the genetic material behind. (Hartwell et al., 2010) This technique was adapted from the initial protocol using gold nano particles for plant cell transformation (Klein et al., 1987, Finer et al., 1990).

Viruses are particles that can be utilized as useful backbone carriers capable of incorporating large amounts of recombinant DNA. Due to the infectious nature of viruses, they are easily able to enter the target cell. The lentivirus contains a vesicular stomatitis virus G glycoprotein responsible for interacting with the phospholipid

component on the cell membrane initiating viral entry by membrane fusion (Burns et al., 1993). Viral entry is not dependant on specific protein receptors, therefore permitting entry into a range of host cells (Burns et al., 1993). The virus is frequently disabled to prevent it from hijacking the host's machinery and replicating uncontrollably. Lenti-viruses, a class of retro viruses, have an added advantage; they contain long terminal repeats (LTRs) that prevent gene silencing. These provide ideal tools for genetic manipulation. (Hartwell et al., 2010)

Gene expression can be altered either by inhibiting expression or be over-expressing it. The modifications will result in a change in the normal cellular process, which can then be analysed using a variety of methods. Over expression is achieved when the gene of interest is cloned into a vector containing a strong constitutively active promoter such as CMV. The vector containing the gene of interest (GOI) in subsequently transformed into the target cell. Due to the nature of the promoter, the mRNA is continuously transcribed which results in an up regulation of the target protein. (Hartwell et al., 2010)

Inhibiting gene expression can be accomplished using gene targeting however, successful recombination occurs at very low efficiencies with a high rate of false positive results. Gene knock downs can also be achieved by utilizing an ingenious strategy harnessing the abilities of RNA interference (RNAi). RNAi is a process initiated by short double stranded RNA molecules (siRNA or miRNA) that are then recognized by the RNA induced silencing complex (RISC). The complex then locates other mRNA strands that contain the siRNA sequence resulting in its

cleavage, implemented by Dicer, and subsequent gene silencing. (Hartwell et al., 2010, Perrimon et al., 2010)

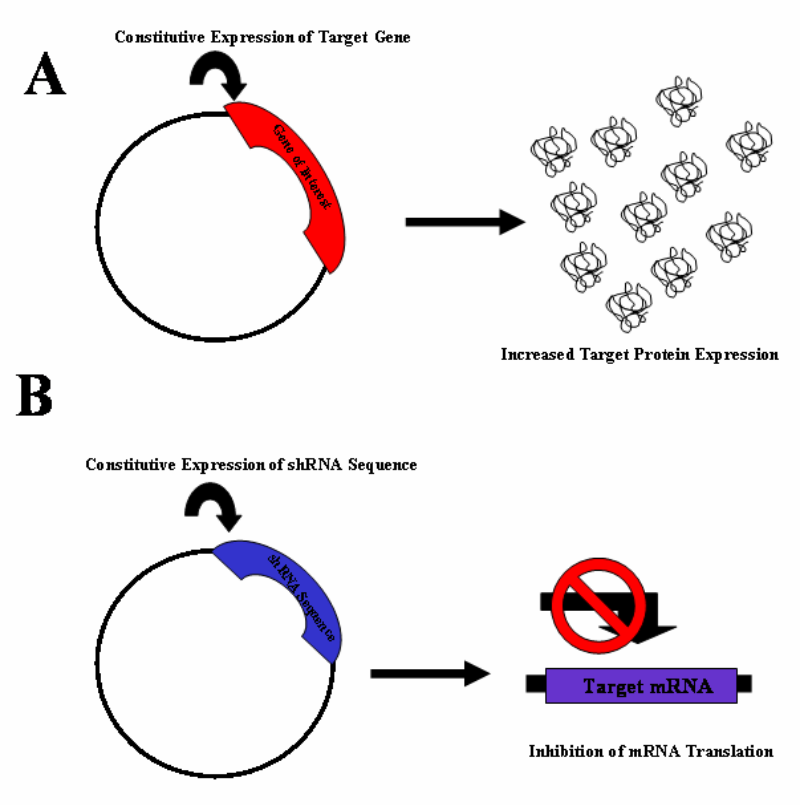


Figure 6.1 – Gene Over Expression and Knock Down
A. The target gene is cloned into a vector containing a strong constitutively active promoter driving the expression of the gene of interest resulting in increased production of the required protein. **B.** Gene knock down is a result of the expression of specific shRNA sequences inducing the degradation of the complimentary mRNA sequence. This inhibits all translational of the target mRNA strand reducing protein levels within the cell.

6.1.3 Challenges Associated with Genetic Modification and hESCs

Genetically modifying hESCs provides a useful tool to study and manipulate a variety of processes. The ability to express or knock down certain genes within hESCs will result in the ability to regulate processes such as directed differentiation, study developmental pathways and alter immunogenicity required for therapeutic applications. Numerous investigations were carried out comparing the various

genetic modification strategies and isolating the most effective techniques. However, the previously employed strategies suffer from low efficiency of success, limited control of gene expression, random mutations and gene silencing. (Strulovici et al., 2007)

Initial methods required to introduce the new genetic material into the hESCs involved chemical transfection or electroporation. Unfortunately, both methods result in low yields due to limited cell viability. These are defined as non-viral forms of modification and their effects are transient, as the introduced gene is not integrated into the target genome. Viral vectors contain glycoproteins present on the surface of their capsid structure, which are able to interact with the surface receptors on the target cells resulting in membrane fusion and viral entry (Burns et al., 1993). The gene or sequence of interest is randomly integrated into the host genome, frequently in multiple locations. Unfortunately, the integrated transgenes are then susceptible to gene silencing and random mutations due to the robust property of hESCs. Lentiviruses, a form of retroviruses, are able to over come this issue as they contain long terminal repeats immune to silencing. Long terminal repeats (LTRs) are 700-1000 nucleotides long and are responsible for inducing and regulating the expression of viral genes vital for integration, replication and packaging (Laimins et al., 1984). Thus, LTRs contain features that inhibit subsequent gene silencing and transposition (Zhang et al., 2010). The 5' LTR acts as an RNA polymerase II promoter whilst the 3' LTR causes transcriptional termination, in addition to providing 'sticky ends' which are integrated into the host genome with the aid of the integrase protein (Takebe et al., 1988). The viral vectors replication gene is usually disabled to prevent

host machinery hijacking and uncontrolled replication. This in turn, allows the safe implementation of lentiviral vectors as tools in genetically manipulating hESCs. (Strulovici et al., 2007)

6.1.4 THE STRATEGY DEVELOPED TO INVESTIGATE THE EFFECT OF SUMOYLATION AND HNF4 α IN LIVING SYSTEMS

Lentiviral vectors are the most effective tools for introducing transgene modifications within hESCs. As a result, we employed the use of the lentiviral vectors, pLVX and pLenti4 to alter the expression of genes involved in the SUMO pathway in conjunction with HNF4 α. The most efficient method of elucidating the effects of SUMO in living systems was to identify changes associated with the over expression and inhibition of the pathway. hESCs were transiently transduced with lentivirus constructs for over expressing SUMO-1 and Ubc9 (E2) (S1OE), SUMO-2 and Ubc9 (S2OE), and Ubc9 was knocked down (U9KD) to reduced SUMOylation. To examine the effect of HNF4 α in our system we created two subsets of cells, one over expressing HNF4 α (H4OE) and the second deficient in HNF4 α expression (H4KD). Gene over expression was implemented by the use of the pLVX vector, whilst pLenti4 was required for siRNA mediated gene silencing. The results discussed in chapter 5 indicate a role of poly-SUMOylation in regulating HNF4 α stability. Thus SUMO-2 over expression constructs should result in the reduction of HNF4 α levels, whilst Ubc9 knock down cells should stabilize HNF4 α levels. No changes in HNF4 α levels should be noted in SUMO-1 over expression cells. In addition, increased hepatic viability is expected in cells over expressing HNF4 α and premature de-differentiation in HNF4 α knock down cells. In order to express

SUMO1/2 and Ubc9 as separate proteins, the 2A sequence was cloned in between the two genes flanked by specific restriction sites. The 2A sequence is a short peptide sequence capable of self-cleavage on translation, resulting in the division of proteins on either end of the peptide (Trichas et al., 2008).

The C3A hepatoma cell line, an important research tool, was initially used to optimize the transfection conditions and investigate the viral integration and expression efficiency. The procedure was then implemented in day 11 differentiating hESCs to investigate the roles of SUMOylation and HNF4 α within our hepatic differentiation system. The technique can then be utilized in primary human hepatocytes, which are regarded as the gold standard control.

6.2 RESULTS

6.2.1 THE CONSTRUCTION OF THE OVER EXPRESSION AND KNOCK DOWN LENTIVIRAL VECTORS

6.2.1.1 THE pLVX OVER EXPRESSION VECTORS

The pLVX lentivector (Clontech, USA) is an HIV-1 based vector with a modified form of the human cytomegalovirus immediate early promoter (P-tight) to allow strong constitutive expression of the transgene, Figure 6.2 A. In also contains the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), which enhances RNA processing and induces transport of the transgene RNA into the nucleus, Figure 6.2 A. This in turn, improves viral titers, the concentration of resulting active packaged virus, and gene expression. pLVX expresses the ampicillin resistance gene allowing for selection in bacteria and the puromycin resistance gene

required for target cell selection, Figure 6.2 A. The presence of ψ sequence induces effective viral packaging carried out in 293T host cells. This vector does not contain its own packaging and replication genes to ensure personal safety when handling the virus in the lab. Additional vectors containing the Pol, Rev, Gag and VSV glycoprotein lentiviral genes are co-transfected with the pLVX vector containing the transgene in 293T cells to promote replication, nuclear translation, core protein expression and viral packaging, respectively. The Gag and Pol genes were expressed under the CMV promoter in the pLP1 plasmid vector also containing the human β-globin intron, which enhances the expression of the gag/pol genes in mammalian cells. The Rev mRNA transcript is generated from the pLP2 plasmid vector under the control of the RSV promoter. The VSV G gene responsible for expressing the glycoprotein required for viral entry is contained within the pLP/VSVG vector in addition to the β-globin intron both under the CMV promoter. The VSV G protein present on the lentivirus permits infection into a number of cell types, including non-dividing cells, thus promoting efficiency of transfection in hESCs and their derivates.

The first step required for creating the lentivectors was to analyze their respective multiple cloning sites and identify unique restriction sites for the insertion of the modified DNA sequence. The SUMO 1 and SUMO 2 DNA sequences were flanked BamHI and NotI restriction sites. A 2A sequence was inserted between the SUMO and Ubc9 genes, flanked by a NotI and a MluI restriction site, to separate the proteins post-translationally. The 2A sequence is a short peptide capable of self cleavage once the DNA sequence has been translated. A MluI and an EcoRI site flanked the Ubc9

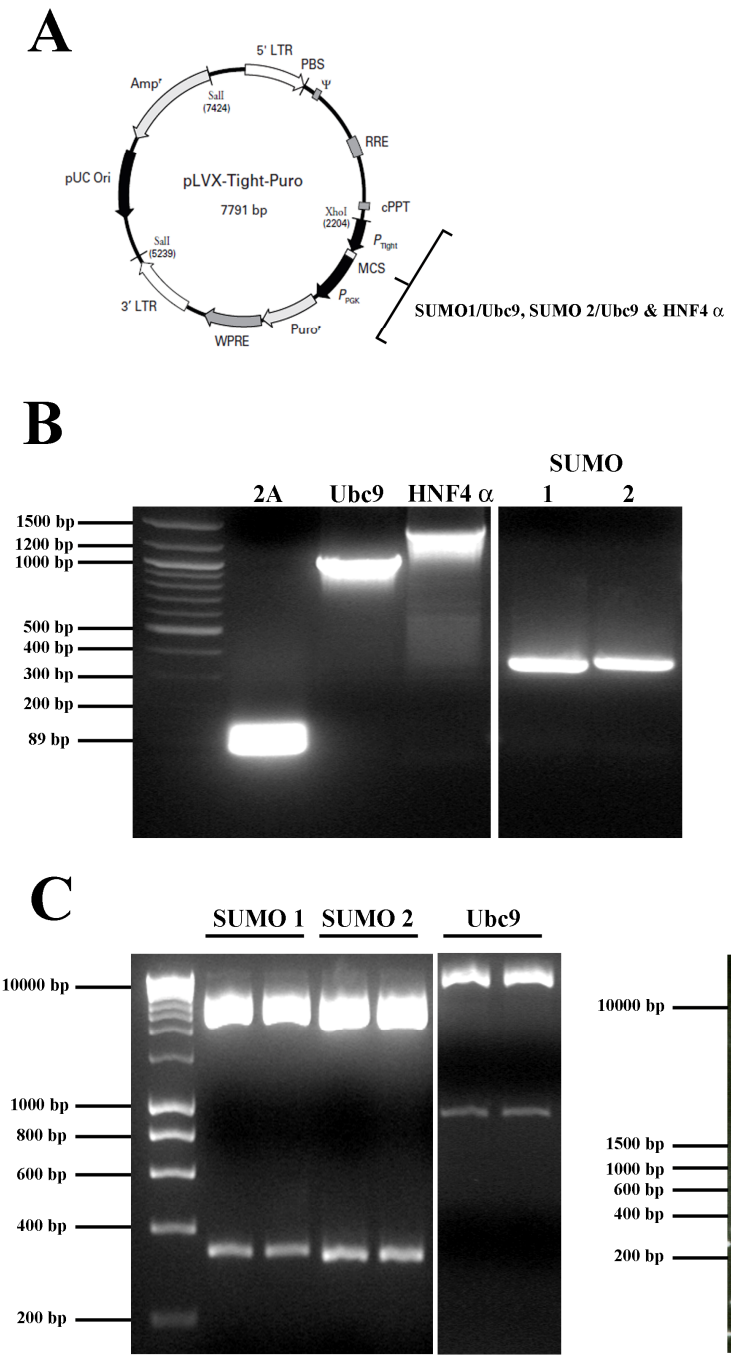


Figure 6.2 – Over Expression Vector Construction

A. The pLVX lentivector is a useful tool for over expressing a transgene. It contains a CMV promoter driving constitutive expression of the inserted gene in conjunction with ampicillin and puromycin resistance genes required for accurate selection. The presence of the ψ sequence signals promotes effective packaging thus providing the necessary machinery to infect a number of various cell types. **B.** The bands depict the accurate expression of the various inserts amplified from their respective image clones. The PCR products were run on an agarose gel and the insert size was verified (base pairs, bp). **C.** A diagnostic digest was carried out to validate the presence of the correct inserts within the PCR 2.1 vector prior to sequences. As noted, all the inserts were present at the correct size, as depicted by the lower bands. The digested PCR 2.1 vector can be identified as the 8000bp band.

sequence. Specific primers for each gene were designed and included the unique restriction sites required for vector construction in conjugation with distinctive protein tags. SUMO-1 and -2 were tagged using the HA sequence and Ubc9 was Myc tagged.

The target genes were amplified from specific image clones using PCR, Figure 6.2 B. The plasmid DNA for the image clones was isolated from bacteria containing the vector using a midi prep kit (QIAGEN, UK). The concentration and quality of the plasmid vector containing the image clone was measured using the Nanodrop. The PCR products were then used in a blunt end ligation reaction (TOPO TA Cloning Kit, Invitrogen) whilst a sample was extracted and run on an agarose gel to assess gene amplification. Ligating the recombinant DNA into the PCR2.1 vector allows for easily manipulation of our transgene and provides a host carrier to carry out analysis before the sequence is cloned into the pLVX vector. The ligated plasmids were transformed into DH5α E-coli cells, which were plated out on agar plates containing ampicillin and incubated overnight. Antibiotic selection ensures the propagation of the host vector as it contains the resistance gene necessary for survival and suggests the successful ligation or recombination of the plasmid vector containing the insert. Clones were selected and cultured to amplify the vector DNA, which was then isolated using a mini prep kit (QIAGEN, UK). This process was repeated for each gene; SUMO1, SUMO2, 2A and Ubc9.

The PCR 2.1 vectors for each transgene were then digested with BamHI and NotI to confirm the SUMO-1 and -2 inserts and MluI and EcoRI to verify the Ubc9

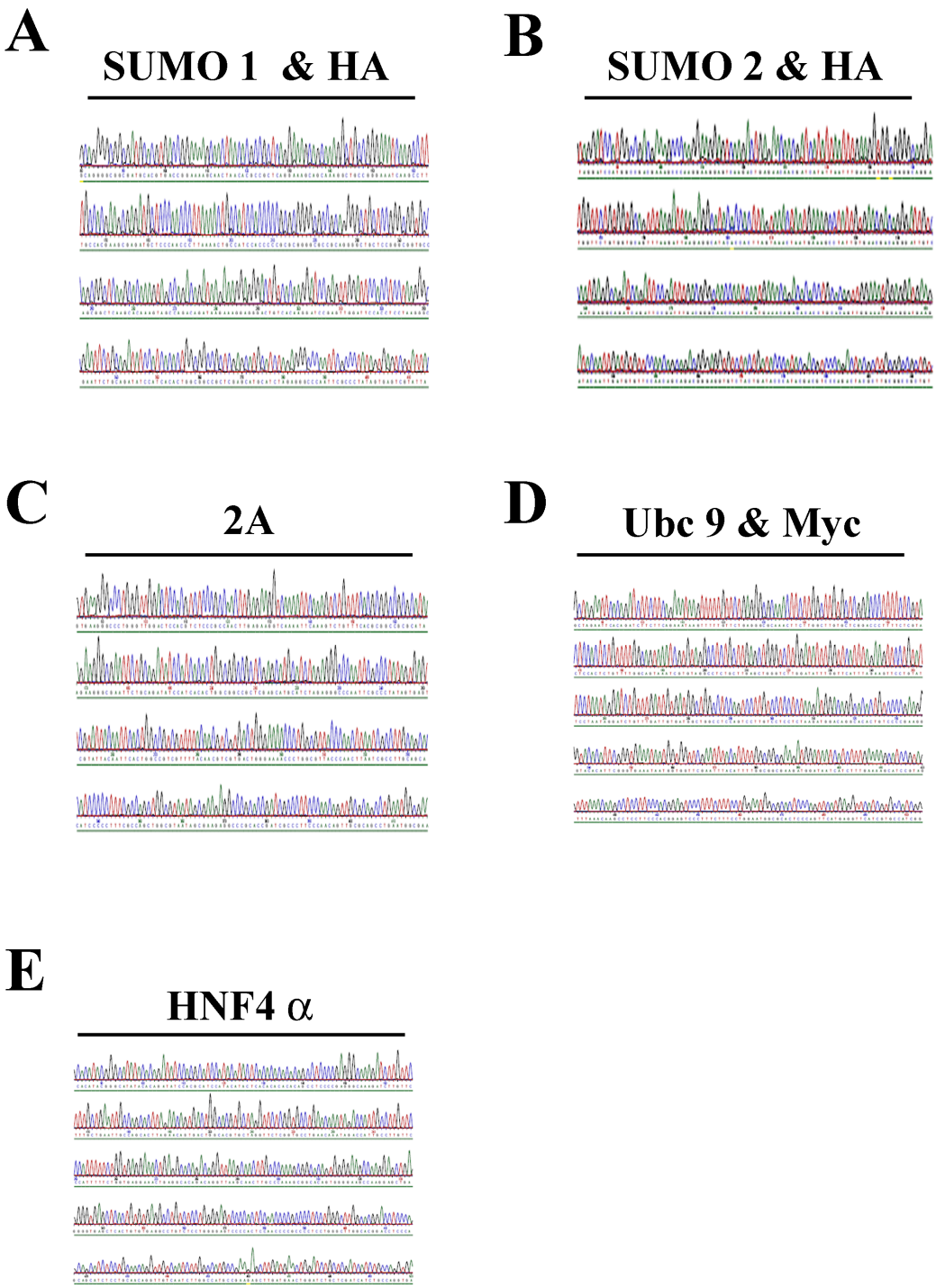


Figure 6.3 – Over Expression Transgene Sequencing
The individual transgene sequences cloned into the PCR 2.1 plasmid vector were sequenced to confirm the amplification accuracy and ensure no mutations had been generated throughout the cloning process. The graphs verify the correct transgene sequences, whereby **A.** confirms the correct SUMO 1 insert tagged with HA, **B.** defines the SUMO 2 and HA tagged insert, **C.** and **D.** validates the 2A and Ubc9 sequence, respectively and **E.** displays the correct HNF4 α sequence.

sequence, respectively, Figure 6.2 C. The digests were run on an agarose gel to validate the presence of the inserts, substantiated by their size (bp), Figure 6.2 C. The correct inserts were verified as noted by the lower bands. The band present at ~ 8000 bp, Figure 6.2 C, defined the digested PCR 2.1 vector, which no longer contained the respective insert. The clones containing the correct insert were then cultured at a larger scale and maxi prepped to isolate the plasmid DNA, which was subsequently sequenced (MWG Eurofins, Germany), Figure 6.3. On sequence verification, the plasmids were digested with the previously mentioned restriction enzymes in conjunction with pLVX vector digestion using BamHI and EcoRI. The inserts were then carefully isolated using the gel extraction and purification technique. After weeks of optimization, a four way ligation was employed and the respective ratios used were as follows; 100ng pLVX: 450ng Ubc9: 250ng SUMO1/2: 50ng 2A. The ligation was incubated at 16°C overnight.

The two ligation reactions were transformed into DH5 α cells, plated out, selected for ampicillin resistance, clones were selected and propagated. The SUMO-1 and -2 pLVX vectors were digested to confirm the presence of the correct insert, Figure 6.4 A. To determine the presence of the full insert, the vector was digested with BamHI and EcoRI. Once the correct sized insert, ~2000 bp, was noted the vectors were purified using a maxi prep kit and sent for sequencing to verify the transgene sequences, Figure 6.4 B-D. The digested pLVX vector is recognized as the band present at ~ 7000 bp. The insert sequences were confirmed, Figure 6.4 B - D, completing the process of constructing the SUMO-1 (B) and SUMO-2 (C) over

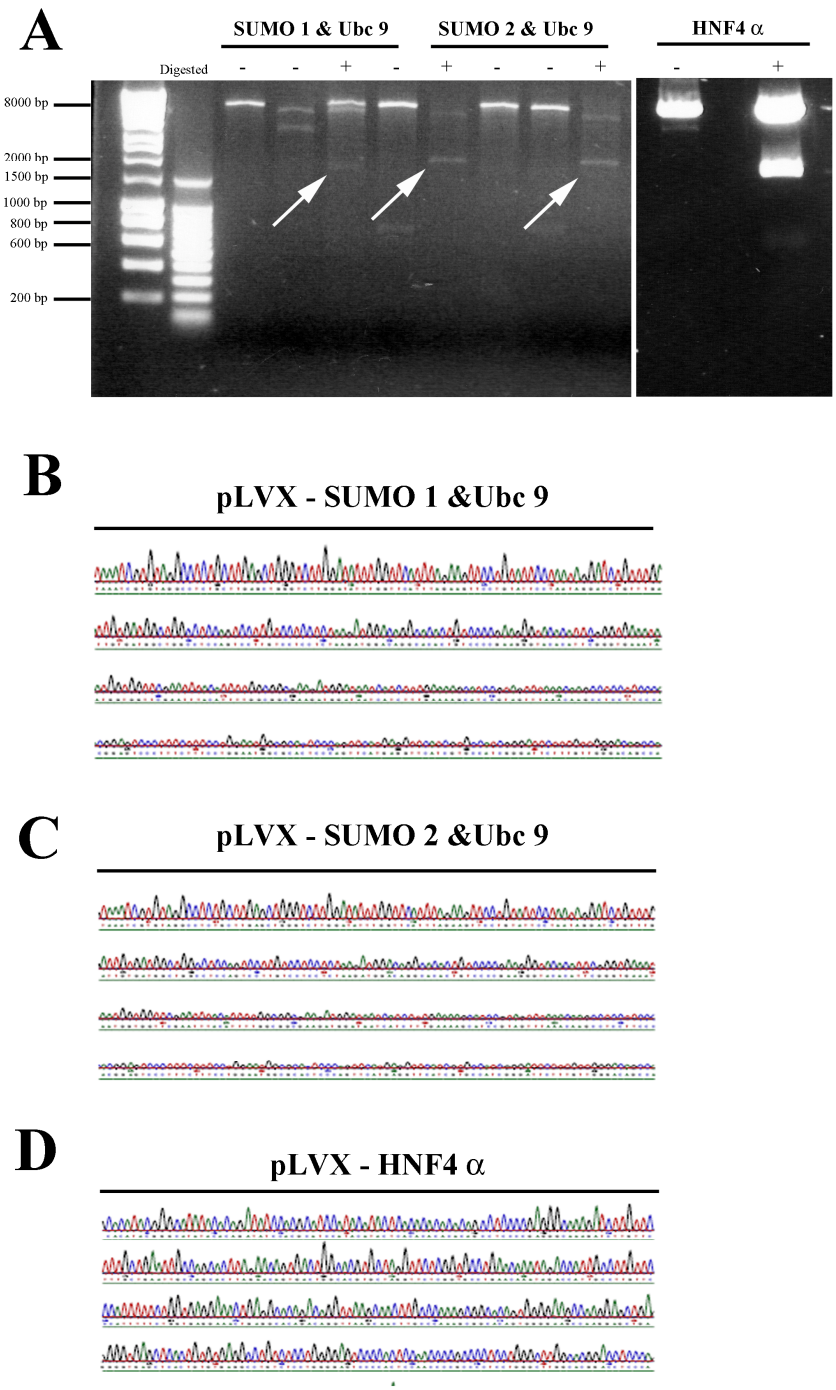


Figure 6.4 – pLVX Over Expression Lentiviral Vector Construction

A. The ligated pLVX vectors containing the transgenes were transformed and propagated in bacterial cells and subsequently purified using a QIAGEN Midi prep kit. The resulting plasmid DNA was digested with BamHI and EcoRI to determine the presence of the full transgene insert. The HNF4 α insert was detected via EcoRI digest. As observed, the bands at the expected sizes were noted and the pLVX vectors were sent for sequencing. **B. C. and D.** verify the correct transgene sequences within the pLVX lentiviral vector.

expression lentiviral vectors. Please note that SUMO-1 and SUMO-2 over expression vectors refer to the full SUMO-2A-Ubc9 insert present in the pLVX vector.

The same techniques were applied for the creation of the HNF4 α over expression lentiviral vector. The full-length sequence of HNF4 α variant three was flanked with EcoRI restriction sites using PCR, Figure 6.2 B. The fragment was cloned into the PCR2.1 plasmid vector and a diagnostic digestion was carried out to confirm the presence of the insert, Figure 6.2 C. Once the insert was observed, the vector was sent for transgene sequencing, Figure 6.3 E. On sequence verification, the insert was EcoRI digested and ligated into the pLVX vector. After the resistant clones were selected and digested to detect the insert (~1500 bp), Figure 6.4 A, the sequence was validated once again to confirm the correct orientation and sequence of the insert. Figure 6.4 D confirms the correct in frame HNF4 α sequence required for gene over expression in our differentiating hESCs.

6.2.1.2 THE PLENTI4 KNOCK DOWN VECTORS

Short hairpin RNA (shRNA) sequences were designed to silence the Ubc9 and HNF4 α genes in hESC derived hepatic endoderm. This technique employs the strategy whereby Ubc9 and HNF4 α mRNA strands are recognized and cleaved in response to the shRNA sequence expression initiated by the pLenti4 lentiviral vector (Invitrogen, UK). The BLOCK-iT™ Inducible H1 Lentiviral RNAi System was applied to generate the Ubc9 and HNF4 α knock down vectors. The knock down system utilizes an entry vector whereby the ligated shRNA sequence recombines into the pLenti4 destination vector, which is subsequently transformed and propagated in Stabl3 E-coli cells, Figure 6.5 A. The entry vector provides and easy and efficient method for

cloning the shRNA sequence into an RNAi specific cassette required for effective shRNA expression resulting in gene knock down. This form of RNAi dependant gene silencing exploits the structure of the shRNA. Once the sequence is expressed, the RNA molecule transforms into a stem-looped structure as the complementary regions at either end anneal. Dicer detects and cleaves these RNA structures forming an siRNA (small interfering RNA) duplex.

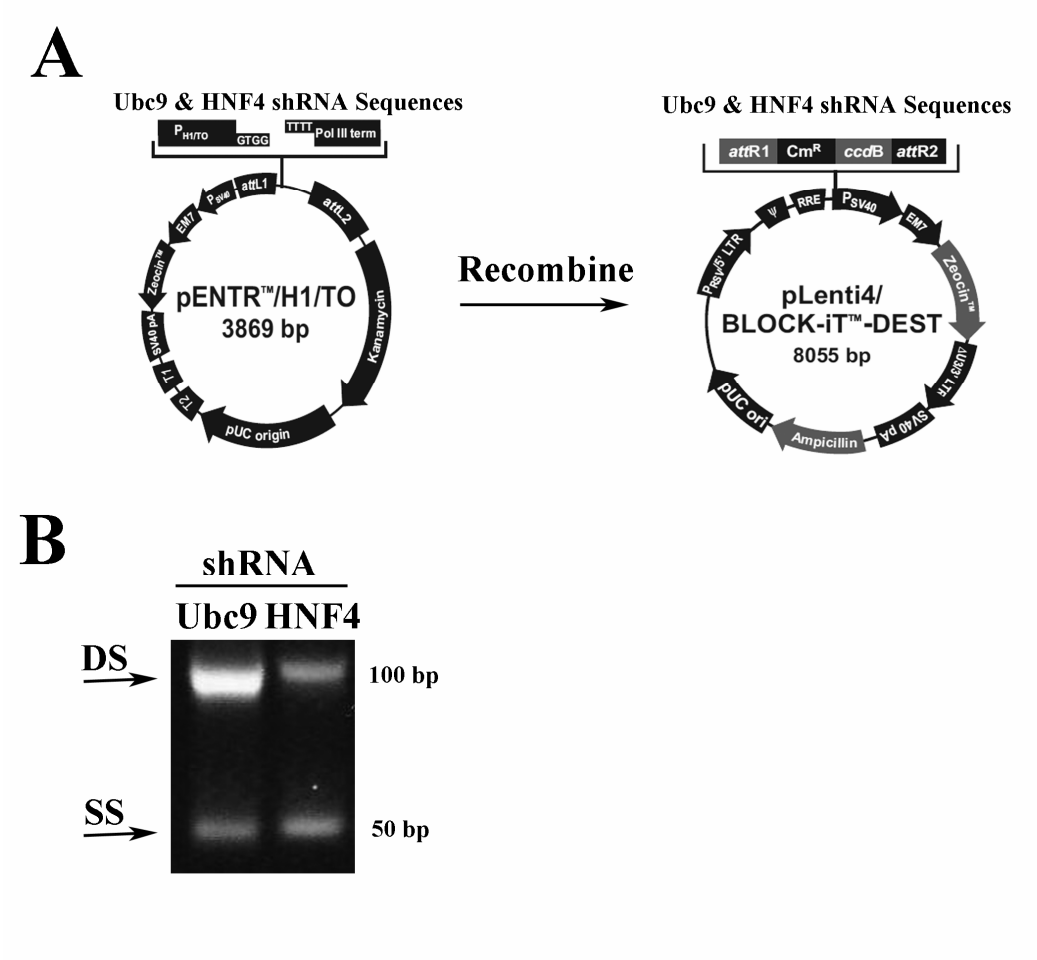


Figure 6.5 – Knock-Down Vector Construction
A. The pENTR™/H1/TO and the pLenti4/BLOCK-iT™-DEST lentivectors are a useful tool for knocking down a specific gene. It contains a human H1 promoter driving expression of shRNA sequence in conjunction with kanamycin (entry vector), ampicillin and zeocin (destination vector) resistance genes required for accurate selection. The presence of the 5' and the ψ sequence in the destination vector promotes effective packaging thus providing the necessary machinery to infect a number of various cell types. Fail-safe mechanisms are also in place to improve the bio-safety of the vector. The entry vector containing the ligated shRNA sequence recombines with pLenti4 to generate the complete knock down vector **B.** The upper bands depict the accurate annealing of the shRNA oligonucleotides ~100 base pairs, suggesting the formation of double stranded DNA complexes (DS arrow) required for successful ligation into the entry vector.

The destination vector, pLenti4/BLOCK-iT™-DEST (pLenti4), contains the Rous sarcoma virus (RSV) enhancer/promoter required for viral mRNA expression in conjunction with the HIV-1 5' LTR necessary for reverse transcription of the viral mRNA, Figure 6.2 A. The ψ sequence and the 5'LTR permit efficient viral packaging, however 5' donor sites and 3' acceptor sites ensures the removal of the ψ sequence and the HIV-1 Rev dependant reverse response element (RRE), involved in unspliced viral mRNA nuclear export, preventing the Rev dependant expression of the shRNA sequence in the target cell Figure 6.5 A. The ΔU3/HIV-1 truncated 3' LTR provides another fail-safe to guarantee the bio-safety of the lentivector, which promotes viral packaging but results in the inactivation of the 5'LTR Figure 6.5 A. The ampicillin resistance gene contributes to selection in E-coli cells whilst the Zeocin resistance (Sh ble) gene allows stable selection in mammalian cell lines. The above elements in addition to the RNAi cassette, containing the human H1 RNA polymerase III based promoter required for shRNA expression, results in the efficient and stable production of RNAi complexes essential for gene silencing.

The Ubc9 and HNF4 α shRNA sequences were designed using the Invitrogen BLOCK-iT™ RNAi Designer. The shRNA complexes consisted of a top strand oligonucleotide complementary to a bottom strand oligonucleotide with 5' and 3' overhangs required for ligation into the entry vector (pENTR™/H1/TO). The shRNA oligonucleotides were annealed as described in the BLOCK-iT™ Invitrogen manual. The reactions were subsequently run on an agarose gel to verify the annealing efficiency, Figure 6.5 B, confirmed by the detection of double stranded DNA sequences (DS arrow). Annealing reactions require an efficiency of 50% or greater

for successful ligation to occur. The annealed reactions were combined with the entry vector and the T4 DNA ligase under optimal ligating conditions. The ligation was incubated at room temperature for two hours to ensure improved colony yields. The entry vectors containing the ligated shRNA sequence were then transformed into One Shot® TOP10 competent E-coli permitting vector propagation and kanamycin dependant selection. Resistant colonies, signifying the presence of the entry vector containing the shRNA insert, were picked and cultured. The vector DNA was then isolated using a midi prep kit (QIAGEN, UK). The transformants were then sequenced to verify the double stranded shRNA oligonucleotides. On sequence confirmation, one clone was selected propagated and the vector DNA was isolated using a maxi prep kit (QIAGEN, UK) to improve DNA retrieval and purity.

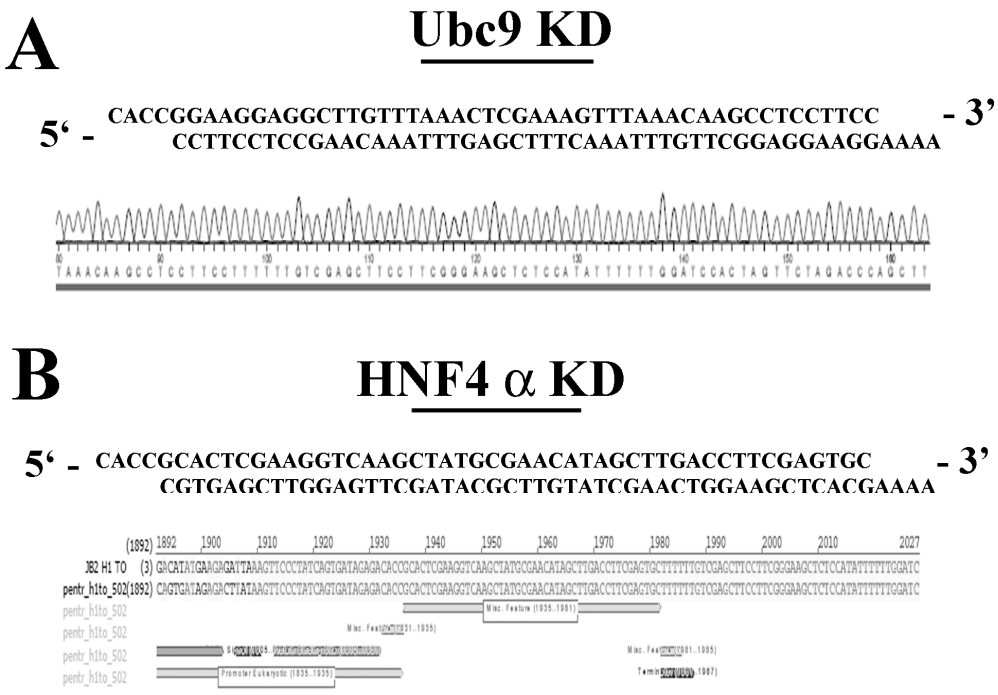


Figure 6.6 – Knock-Down Lentiviral Vector Sequencing
The sequencing graphs confirm the accurate annealing and recombination of the Ubc9 **(A)** and HNF4 α **(B)** shRNA sequences in the pLenti4 destination vector required for gene knockdowns.

The LR recombination reaction, occurring between the attL and attR sites, was set up using the entry vector positive for the correct shRNA sequence and the pLenti4 destination vector and left for 16 hours to promote higher colony yields. Proteinase K was used to arrest the reaction, which was then transformed and ampicillin selected in Stabl3™ competent E-coli. Resistant clones were chosen and propagated, followed by vector DNA isolation using a maxi prep kit (QIAGEN, UK). The resulting vector DNA was sent for sequencing using the H1 forward and V5 reverse primers to verify the insert. Figure 6.6 confirms the correct shRNA sequences of the Ubc9 and HNF4 α genes within the pLenti4 destination vector. The knock down vectors are ready for generating Ubc9 and HNF4 α knock down cells.

6.2.2 GENERATING INFECTIOUS OVER EXPRESSION AND KNOCK DOWN LENTIVIRAL VECTORS

6.2.2.1 VIRAL PACKAGING

The next stage of pathway modification required the packaging of the three over expression vectors; SUMO1/Ubc9, SUMO 2/Ubc9 and HNF4 α in conjunction with the knock down vectors; Ubc9 and HNF4 α. A self ligated pLVX vector was used as a negative control for the over expression vectors and a scrambled shRNA sequence was used as a knock down negative control. This ensures that any observable changes are a result of the over expression or knock down and not due to the lentiviral transfection process. Seven lentiviral vectors were used for this investigation. To simplify the nomenclature of each lentiviral vectors the following acronyms were used; S1OE, S2OE and H4OE refer to the SUMO1, SUMO2 and

HNF4 α over expression vectors. U9KD and H4KD define the Ubc9 and HNF4 α knock down vectors. CT is used to identify the respective control lentiviral vectors.

The lentiviral vectors (LVs) were initially incubated with the ViraPower™ packaging mix in Opti-MEM® media, which promotes the formation of the DNA-lipofectamine® 2000 complexes. The packaging mix contains the pLP1, pLP2, pLP/VSVG plasmids in optimal concentrations to induce high efficiency viral packaging promoting mammalian cell infection. Lipofectamine ® 2000 was diluted in Opti-MEM® in a separate sterile 5ml tube. The LVs and the packaging mix were combined with lipofectamine®2000 and were incubated for 20 minutes at room temperature. This provides ample time for the DNA-lipofectamine®2000 complexes to form. This form of transfection utilizes the cationic lipid based properties of lipofectamine®2000 to alter the cellular plasma membrane and allow nucleic acid transfer into the cell. The DNA-lipid complexes are then combined with resuspended 293FT cells and incubated overnight at 37°C. The human 293FT cell line is a derivative of the human embryonic kidney 293 cell lines where the SV40 large T antigen is constitutively expressed. 293FT cells are ideal for efficient lentiviral production as they can be easily transfected, they retain robust viability in culture and the SV40 antigen ensures high levels of viral protein expression vital for capsule formation. The media is replenished 24 hours post transfection and the supernatants are harvested 48-72 hours later, however no significant differences in viral yields were noted between these two time points. The supernatants were purified using centrifugation and filtration, and stored at -80°C for later use.

6.2.2.2 LENTIVIRAL VECTOR TITERS

Titrating lentiviral stocks provides a measure of the virus particle concentration present in the supernatant, which assesses the efficiency of target cell transfection. In addition, the viral titer can also be used to regulate the number of integrated copies of the vector in the host genome as well as the generation of reproducible results. The viral titer for the seven LVs was calculated by transfecting 293FT cells with 10-fold serial dilutions of the viral stock, 1ml starting volume, in the presence of Polybrene® (6µg/ml) (hexadimethrine bromide), a transduction enhancer. The 293FT cells were then selected for 12-14 days using the appropriate antibiotic, puromycin for the over expression vectors and Zeocin® for knock down LVs. The remaining surviving colonies were stained with crystal violet and subsequently counted.

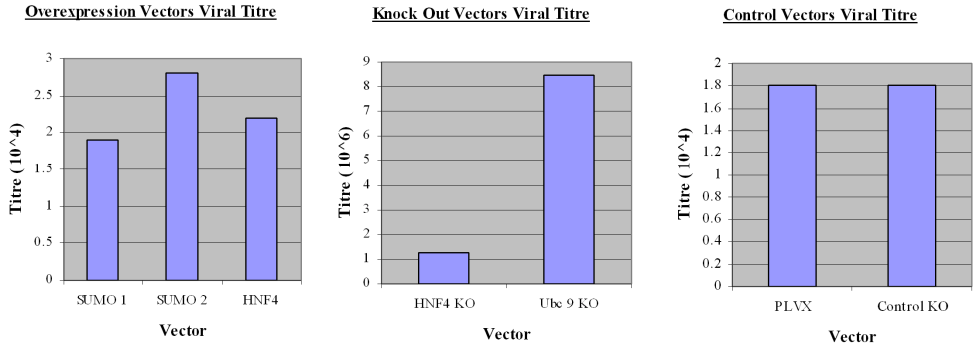


Figure 6.7 – Over Expression and Knock Down Lentiviral Stock Titers
Viral titers provide an effective method to measure transfection efficiency, regulate the number of integrated copies of the vector within the host genome and generate reproducible results. Viral titers were calculated by transfecting 293FT cells with the respective viral supernatant and were then selected for 12-14 days. The remaining colonies were stained with crystal violet and counted. The final titers for S1, S2 and H4 over expression stocks were calculated to be 1.9, 2.8 and 2.2 (x 10^4) TU/ml, respectively. The U9 and H4 knock down vectors were measured to be 8.5 and 1.3 (x 10^5) TU/ml, respectively, with the control vectors titers found to be 1.8(x 10^4) TU/ml.

Expected viral titers range from 5 x 10^5 to 2 x 10^7 transducing units per ml (TU/ml). The over expression vector viral titers for S1, S2 and H4 were calculated to be 1.9, 2.8 and 2.2 (x 10^4) TU/ml, respectively, Figure 6.8. The U9 and H4 knock down

vectors resulted in higher titers measured to be 8.5 and 1.3 (x 10^5) TU/ml, respectively, with the control vectors titers found to be 1.8(x 10^4) TU/ml, Figure 6.7. Despite the lower than expected levels of the LV titers, the viral stocks are sufficient for effective transfection as an increased volume will be added per reaction to compensate for the reduced concentration.

6.2.2.3 ANTIBIOTIC KILL CURVES

Antibiotic kill curve experiments were carried out simultaneously with the viral packaging process. Kill curves are generated to assess the optimum antibiotic concentration used for target cell selection.

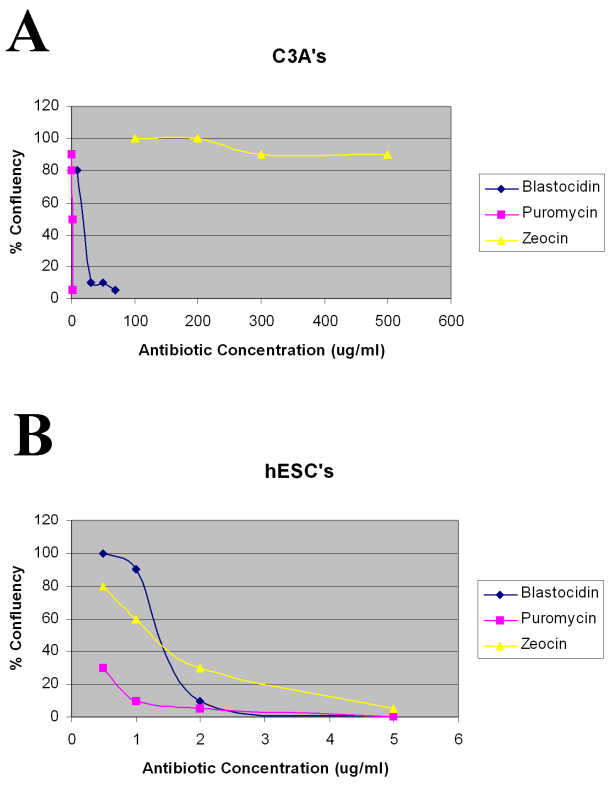


Figure 6.8 – Antibiotic Kill Curves
C3A cells (**A**) and hESCs (**B**) were incubated with varied concentrations of blastocydin, puromycin and Zeocin®. A kill curve was generated to assess the optimum concentration of antibiotic selection, which is defined by the death of 50-70% of the cells. The optimum antibiotic concentrations for C3A cells were calculated to be 20µg/ml, 1.5µg/ml and resistant, respectively. hESCs proved to be more sensitive due to their fragile nature in response to stressed environments. The optimum concentrations were found to be .5 µg/ml, 0.2 µg/ml and 0.75 µg/ml, respectively.

Optimum antibiotic concentrations vary widely between different cell lines and culture conditions; as such they should be assessed before each set of transfection experiments. Optimum antibiotic concentrations are defined when 50%-70% of the confluent cells lift and die. Positive selection ensures that the surviving cells efficiently express the lentiviral vector containing the sequence of interest in addition to the antibiotic resistance gene.

hESCs and C3A cell lines were incubated with various concentrations of the blastocydin, puromycin and Zeocin® antibiotics, ranges applied were generated using data from previous investigations. Figure 6.8 A displays the kill curves required for C3A cell selection. The optimum antibiotic concentrations blastocydin and puromycin were found to be 20µg/ml and 1.5µg/ml, respectively. Interestingly, the C3A's seemed to be resistant to Zeocin® with only 10% cell death at a concentration of 1mg/ml, suggesting the inefficiency of this form of selection. hESCs, on the other hand, were found to be more sensitive to antibiotic selection than C3A's. Figure 6.8 B displays the kill curves for the mentioned antibiotics, whereby the optimum concentrations of blastocydin, puromycin and Zeocin® noted were 1.5 µg/ml, 0.2 µg/ml and 0.75 µg/ml, respectively. The increased sensitivity of hESCs can be expected, as the cells are more fragile with reduced cell viability in stress-induced conditions when compared to the more robust nature of the cancer C3A cell line.

6.2.3 VERIFICATION OF LENTIVIRAL GENE KNOCK DOWN AND OVER EXPRESSION

Prior to the large-scale pathway manipulation in primary human hepatocytes, the C3A cell line and hepatic derivates of hESCs, we set up a preliminary transfection to confirm the vectors ability to knock down and over express the required genes. C3A cells were the ideal cell line of choice due to their robust method of culture, high transfection efficiency and rapid cellular replication. 1ml of each viral vector stock was incubated with C3A cells at 20-30% confluency for approximately 24 hours, achieving a multiplicity of infection (MOI) of approximately 5. It should be noted that C3A cells were examined and confirmed to express the SUMO pathway machinery including SUMO 1 (10 kDa), SUMO 2 (10 kDa), Ubc9 (18 kDa) and SENP 7 (110 kDa), using western blotting, Figure 6.9. β-actin was used as a protein loading control.

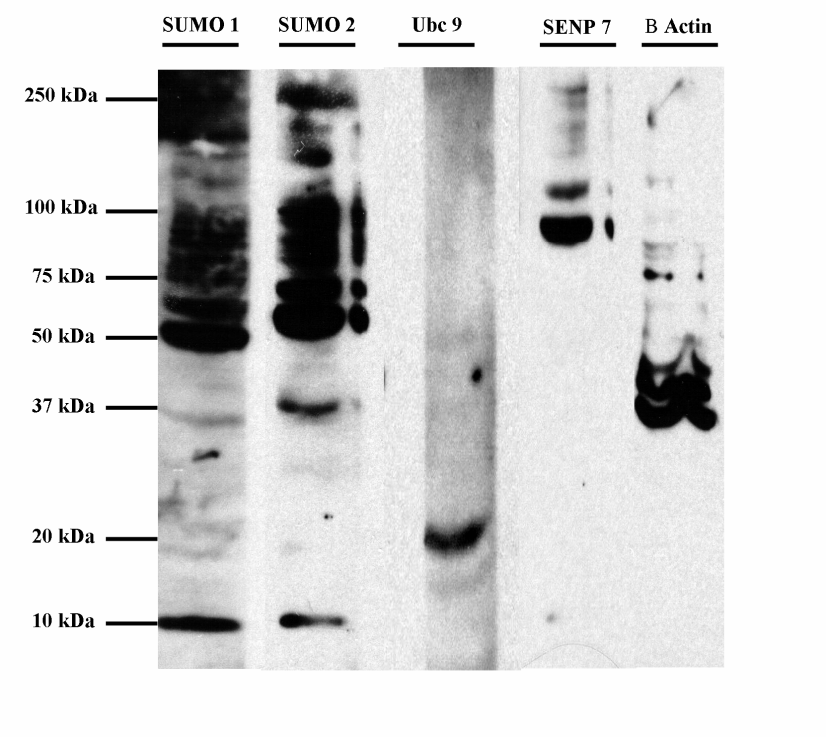


Figure 6.9 – Identification of SUMO Machinery within C3A Cells
The western blots confirm the expression of the required SUMO machinery analysed from C3A cell extracts. Specifically, SUMO1/2, Ubc9 and SENP 7 were sufficiently expressed within the cells. β-actin was used as a loading control.

The C3A cells were cultured for another 48 hours, after which their protein extracts were collected and analysed for gene knock down and over expression using western blotting, Figure 6.10.

Figure 6.10 A displays the knock down effects of the HNF4 α and Ubc9 LVs. All changes in expression were calculated using densitometry analysis on Image J. The HNF4 α LV resulted in 100% successful knock down in the C3A cells, whilst Ubc9 depicted gene knock down by a lesser extent, 50% decrease (arrows). Densitometry analysis displays greater than 5 fold decrease in the expression of HNF4 α and less than two fold decrease in Ubc9 within the knock down cells. With respect to the over expression vectors, Figure 6.11; SUMO1, SUMO2, Ubc9 and HNF4 α demonstrated a significant increase in protein expression; 1.2, 1.7, 1.4 and 2.2 fold increase between the cells transformed with the over expression LVs compared to the empty vector control, Figure 6.11. Knock down cells display a reduction in endogenous expression whilst over expression samples define an increase in expression.

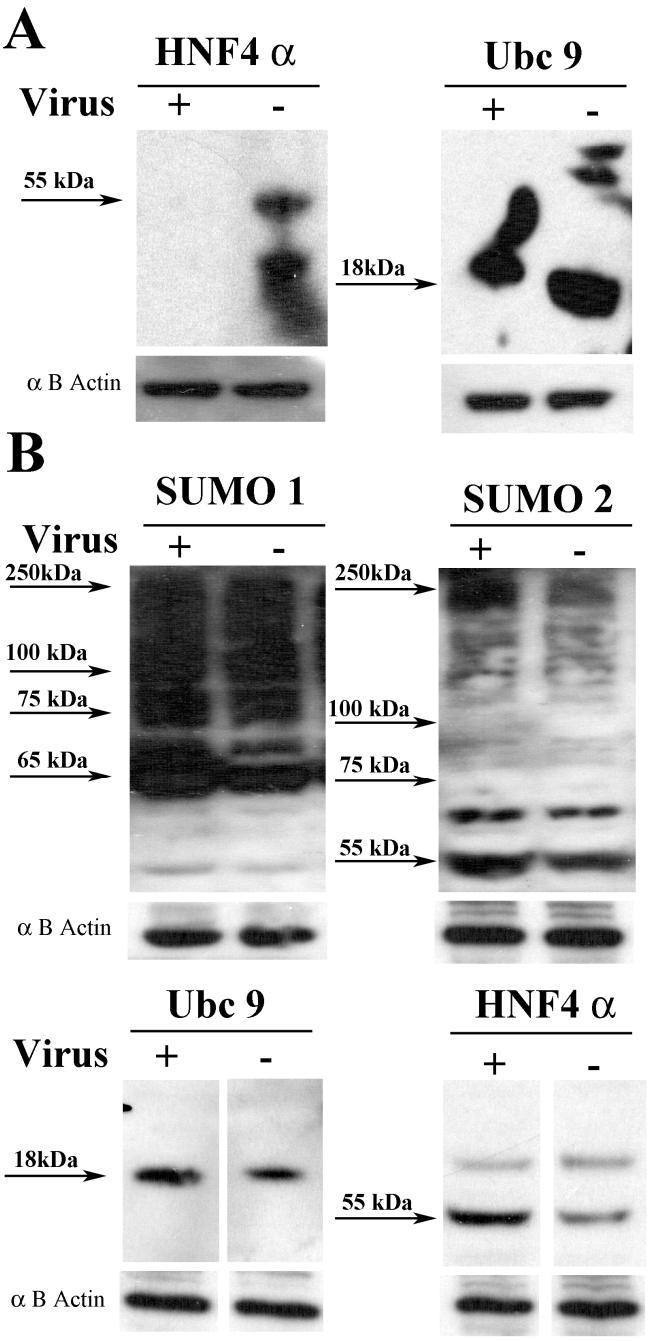


Figure 6.10 – Verification of Lentiviral Gene Knock Down and Over Expression, Western Blotting

A. The western blots show 100% HNF4 α knock down whilst Ubc9 displayed 50% gene silencing, ideal for cell survival and significant enough to detect a notable change within the cell (arrows). **B.** The over expression vectors were able to significantly enhance protein expression of the SUMO1/2, Ubc9 and HNF4 α genes, demonstrated by the darker bands (arrows). Overall, this data has confirmed the ability of the LVs to effectively transform, knock down and over expression the genes of interest. The manipulation of the gene expression pattern should create an observable difference within each living system.

Despite the reduced efficiency of Ubc9 knock down, the resulting advantage is improved cell survival as previous investigations stated in chapter 5 suggest the requirement of Ubc9 expression for healthy cell growth. β actin was used as a loading control. This data confirms the ability of the LVs to efficiently transfect mammalian cells lines resulting in a high level of gene knock down and protein expression. The genetic modification of these pathways should create a significant and observable change within the living systems.

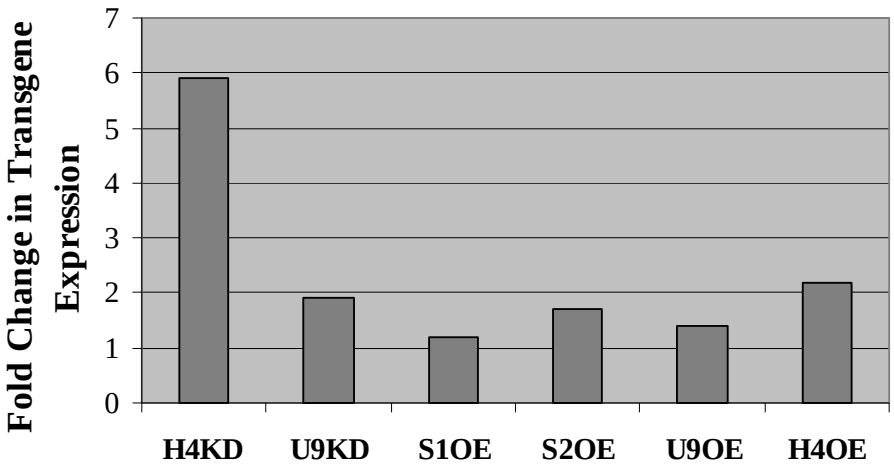


Figure 6.11 – Verification of Lentiviral Gene Knock Down and Over Expression, Graphical Representation
The graph displays the densitometry analysis of the change in expression of the respective transgenes. Knock downs refer to a fold decrease whilst over expressions define a fold increase. As noted, HNF4 α knock down was the most efficient, followed by a less than 2 fold decrease in Ubc9 expression. SUMO 1/2, Ubc9 and HNF4 α averaged a 1.5 increase in expression when compared to the controls.

6.2.4 INVESTIGATING THE EFFECT OF SUMO AND HNF4 α PATHWAY MANIPULATION *IN VITRO*

Genetically modifying the expression patterns of the SUMO and HNF4 α pathway in differentiating hESCs may result in an identifiable phenotypic change relevant to hepatic differentiation, viability and function. The transient modification experiments were carried out, whereby the differentiating hESCs were separated into three

groups; over expression, knockdowns and control groups. The differentiating hESCs were infected with virus at an MOI of one. hESCs were differentiated into HE using the standard differentiation protocol and were subsequently transfected with viral supernatant at day 11, hepatic specification and maturation stage. Day 11 was selected to be an effective time point for pathway manipulation, as this stage defines the beginning of hepatic maturation. As such, altering the signalling pathways at this stage should affect the end stage hepatic phenotype; morphologically, transcriptionally or functionally. The resulting HE were cultured for a further 3 days (Day 20) to monitor the effects on hepatic de-differentiation and viability. Protein and RNA extracts were collected at day 13, 15, 17 and 20, in addition to functional analysis of the hepatocytes on days 17 and 20.

Effective transfection and the subsequent effects were characterized using a range of accepted techniques. Vector integration into the host genome was verified by amplification of the sample mRNA through PCR using primers that exclusively recognize the foreign sequences of the lentiviral vector. The viral supernatant was removed 24 hours after infection, and the cells were washed several times to remove all non-integrated viral particles. This ensured that viral sequence detected through PCR was in fact an integrated vector within the host genome. The PCR utilized primers specific to the 5' and 3' long terminal repeats within the lentiviral vectors and as such only detect integrated viral particles within the host genome. Western blotting confirmed the down stream translational effects of gene knockdown and over expression, with respect to the target proteins and other hepatic markers defining maturity, function and viability.

6.2.4.1 Assessing Viral Integration into the Host Genome

PCR validated the successful integration of all viral vectors into both day 17 and 20 hESC derived HE, denoted by bands present at 100bp for the knockdown lentiviral vectors (A) and 1200 bp for the over expression lentiviral vectors (B), Figure 6.12. However, slightly reduced pLVX copies were integrated into the host genome as indicated by the faint 1200 bp band at day 17, Figure 6.12 B. The bands present at 50 base pairs in both the knockdown and over expression groups depict primer dimers.

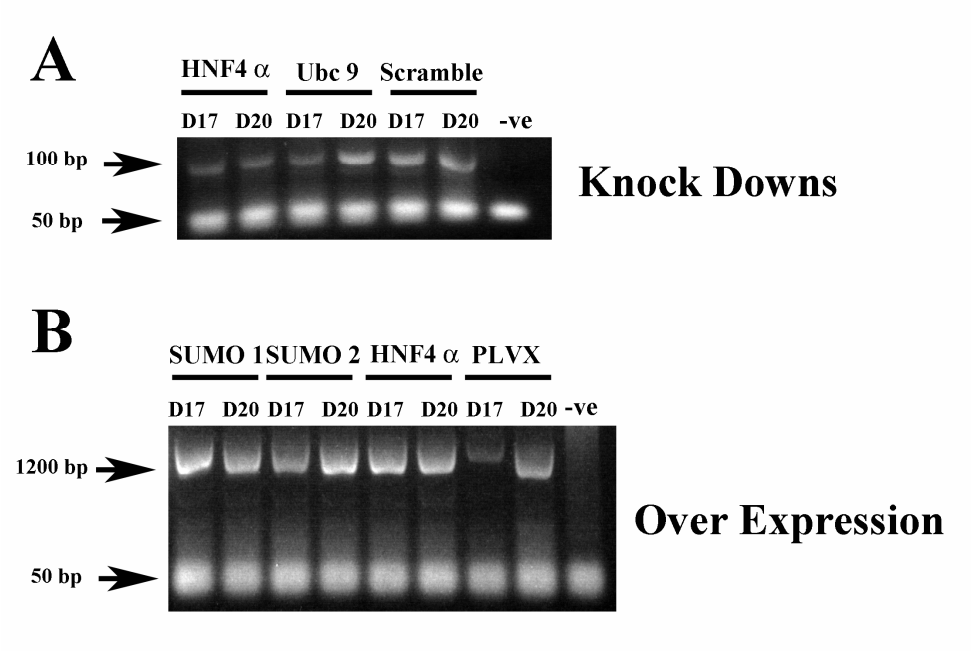


Figure 6.12 – Assessing Lentiviral Vector Integration into the hESC Derived HE Genome
A. Bands present at 100bp (arrow) suggests successful integration of the knockdown vectors into the host genome. **B.** The 1200 bp bands denote the presence of the over expression vectors in the host genome, observed in all four lentiviral vectors (arrow). Bands at 50 bp are due to primer dimers (arrow). A negative no template control was used to remove false positives. This data confirms the successful integration of all seven LVs into the host genome, signifying efficient transfection.

The effective transfection and integration of the lentiviral vectors into the hESC derived HE genome indicates optimal MOI and infection conditions, despite the sub-optimal confluency of the cells during transfection (~70%).

6.2.4.2 Investigating the Translational Effects of Altering Gene Expression

In spite of the successful HNF4 α LVs integration into the host genome, there were no significant differences in HNF4 α expression between the knockdown and over expression samples when compared to the control, Figure 6.13 A. This suggests the unreliable nature of LV transfection, as only a collection of samples were successfully infected. The differences between the transcriptional and translational results could also be attributed to subsequent gene silencing in the hESC derived HE. Consequently, the expression levels of the mature markers, Albumin and Cyp 3A4, and de-differentiation markers, E-cadherin and vimentin remained constant, Figure 6.13 A. A similar pattern was noted in the SUMOylation over expression and knock down cells where no observable difference could be noted between the samples, Figure 6.13 B. As such, no defined conclusions could be drawn from this investigation and further refinement of the transfection technique is required.

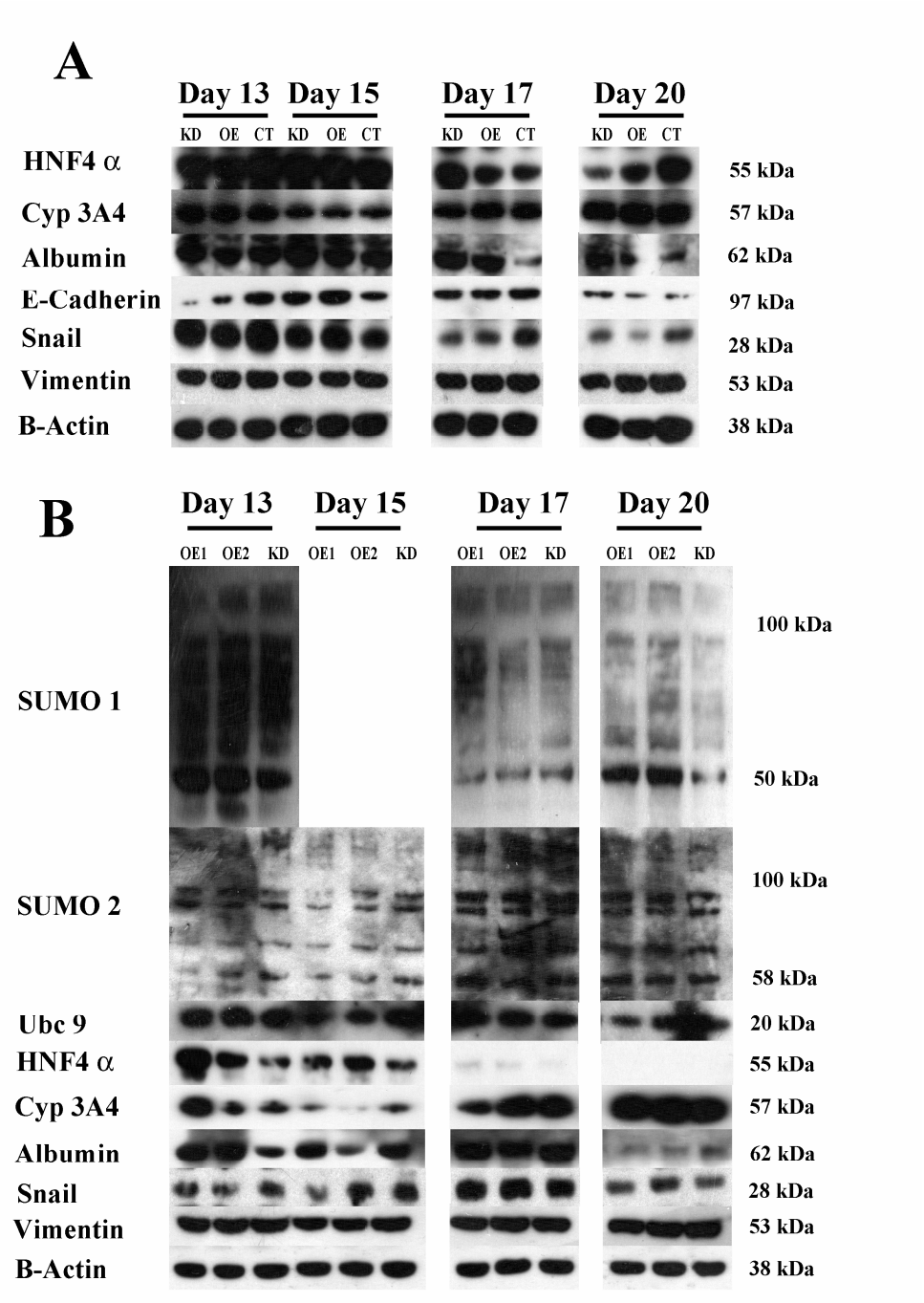


Figure 6.13– Investigating Protein Expression Post-Transfection

A. The blots displayed refer to the sample of cells targeted for HNF4 α over expression and knockdown. No noticeable changes were observed in the levels of HNF4 α expression, with no subsequent changes in maturity and viability marker expression. E-cadherin expression peaked at day 17, suggesting a stable state of HE. Contrary, to previous results, the levels of HNF4 α and albumin only decreased slightly as the cells matured and began de-differentiating. **B.** HE transfected with SUMO over expression and knockdown vectors displayed the expected expression pattern, as noted in previous experiments. The levels of SUMO modified proteins decrease as differentiation proceeds, followed by the decrease in HNF4 α and albumin expression. SUMO 1 and 2 was successfully over expressed in day 20 samples; however, no further change in other factors was noted. Therefore, no firm conclusions can be drawn from this investigation.

6.2.4.3 The Effect of HNF4 α Knock Down and Over Expression in Apoptosis

Caspase 3 is a member of the cysteine- aspartic acid protease family and is responsible for interacting with caspase 8 and 9 resulting in apoptosis (Alnemri et al., 1996). Caspase 3 is activated when triggered by extrinsic death ligands or intrinsic mechanisms, for example mitochondrial stimulation (Salvesen, 2002). Therefore, caspase 3 is a good indicator of apoptosis. Caspase 3 activity was measured in a manner similar to Cyp 3A4. A caspase 3 substrate conjugated to luciferin was incubated with C3A cells transformed with lentiviral vector. C3A cells transfected with pLVX were used as a base line control.

Apoptosis was measured in day 20 hESC derived HE transfected with knock down the HNF4 α LV. Caspase 3, an acknowledged method for identifying apoptosis, activity was measured in the knock out cells and in the vector control. The reduction in HNF4 α expression resulted in a 1.5 fold increase in caspase 3 activity (2.3 million RLU/mg of protein/ hr +/- 0.18, p=0.003), Figure 6.14. The up regulation of caspase 3 in HNF4 α knockdown HE may suggest a role of HNF4 α in sustaining hepatic viability, in vivo, despite the differences observed in the western blots.

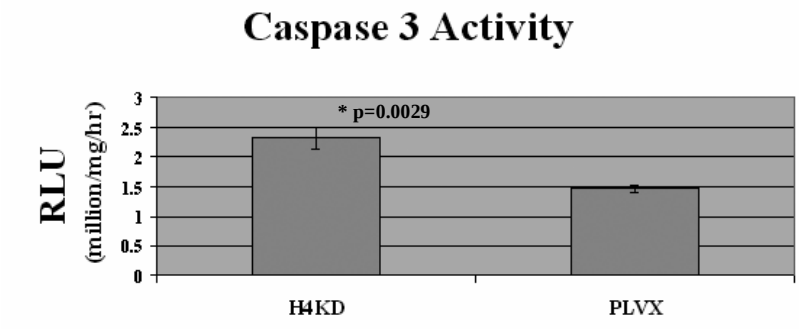


Figure 6.14 – The Effect of HNF4 α Knock Down in HE Apoptosis
Caspase 3 activity was used to measure the rate of apoptosis within hESC derived HE transfected with the HNF4 α knock down vector and the pLVX vector, used as the control. Caspase 3 levels were found to be 1.5 fold higher in HNF4 α knock down HE when compared to the control (2.3 million RLU/mg of protein/ hr +/- 0.18, p=0.003). This data suggests that HNF4 α plays a role in maintaining HE viability in culture.

In addition, caspase 3 activity was measured in C3A cells over expressing HNF4 α. Caspase 3 activity was significantly reduced in cells containing the HNF4 α over expression vector when compared to the pLVX infected cells (p=0.0006), Figure 6.15. Caspase 3 levels in C3As over expressing HNF4 α were 2.5 fold lower than the control cells; respective values were calculated to be 0.34 (+/- 0.02) and 0.85 (+/- 0.08) RLU (million/mg of protein/hr), signifying the possible role of HNF4 α in reducing cell apoptosis and therefore improve cell viability.

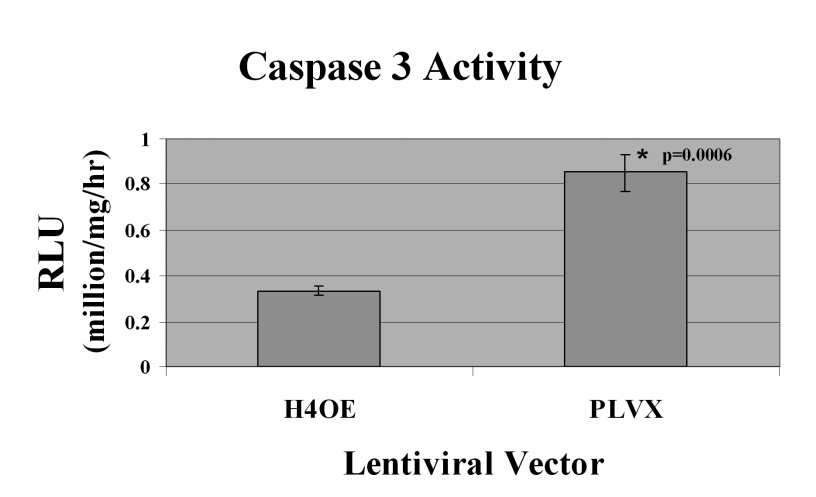


Figure 6.15 – The Effect of HNF4 α Over Expression in C3A Cell Apoptosis
Caspase 3 activity was used to measure the rate of apoptosis within C3A cells transfected with the HNF4 α over expression vector and the pLVX vector, used as the control. Caspase 3 levels were found to be 2.5 fold lower in cells over expressing HNF4 α when compared to control cells (p=0.0006). The caspase 3 values for C3As expressing HNF4 α and pLVX were 0.34 (+/- 0.02) and 0.85 (+/- 0.08) RLU (million/ mg of protein/ hr), respectively.

Overall, the caspase 3 data in these two cell types suggest that HNF4 α may in fact play a direct role in maintaining hepatic viability, whereby HNF4 α knock down cells displayed an increase in apoptotic events and cells with HNF4 α over expression demonstrated a reduced number of cells entering apoptosis within the population. However, further investigation with a more refined system is required before conclusive data could be drawn.

6.3 DISCUSSION

Real time investigation of important pathways in living systems provides significant insight into the mechanisms in action and their subsequent effects. Therefore, the use of lentiviral vectors to over express and knockdown specific genes is an ideal technique utilized for effective pathway manipulation. hESC differentiation into hepatic endoderm has been shown to be efficient however; the hepatocytes are not viable for long term culture. This observation is also noted in freshly isolated primary hepatocytes cultured in vitro. Previous investigations in conjunction with literature reviews suggest a possible role for HNF4 α in maintaining hepatic stability in culture. Our data has also shown that SUMOylation plays a role in maintaining HNF4 α stability in both vitro and vivo. As such, gene expression of HNF4 α and SUMO pathway components were optimized in the hepatoma C3A cell line and subsequently altered in hESC derived HE. The technique can eventually be used in primary human hepatocytes, the gold standard model in research.

The lentiviral vectors were packaged, titered and a trial transfection was set up in C3A cells to investigate transfection and expression efficiency. All vectors were successfully able to infect, integrate and significantly express the required sequences, as observed in Figure 6.10. Therefore, the transfection procedure was effectively optimized in the C3A cell line. The next stage of the experiment was to investigate the effect of HNF4 α and SUMO modification in hESC derived HE. The combined results highlight the disadvantages associated with lenti viral transfection within hESCs. Lentiviral vectors were demonstrated to integrate into the differentiating hESC genome; however, no subsequent protein expression differences were noted.

Therefore, no defined conclusions could be drawn. Caspase 3 activity was associated with changes in HNF4 α expression, indicating that HNF4 α may play a direct role in maintain hepatic viability. However, reproducing the data in a more refined model is required before drawing definitive conclusions. Therefore, optimization in C3A cell lines was not sufficient in generating reproducible results in differentiating hESCs. Thus, further investigation is required to generate suitable transfection conditions for successful manipulation of expression pathways within hESCs.

There are a number of factors that have been attributed to affecting the transformation efficiency of lentivirus. The optimal multiplicity of infection (MOI) seems to be different between various cell lines and types, in addition to the confluency of the cells at the point on transfection. High levels of viral infection and integration have been demonstrated when transfecting cells at 20-30% confluency. Unfortunately, due to our differentiation model, this was not possible as passaging hepatic endoderm during differentiation would result in significant cell death and loss of the hepatic phenotype within minimal replating efficiency. Studies have shown that introducing the hepatitis B virus (HBV) enhancers and the post-transcriptional regulatory element into a lentiviral vector significantly improves hepatocyte transfection efficiency (Seppen et al., 2002). In addition, Selden and colleagues have demonstrated that the use of growth factors, specifically hepatocyte growth factor (HGF) and epidermal growth factor (EGF), resulted in the improvement of the expression level of the transgene and increased levels of expression per cell, signifying a greater number of integrated viral copies (Selden et al., 2007). Viral handling has also been found to affect transfection ability. Long-

term storage of viral supernatants at -80°C in addition to repetitive freezing and thawing results in a significant reduction in transfection efficiency. Overall, improvement of transient transfection in differentiating hESCs may require increasing the MOI, using freshly generated viral supernatant and creating a more stable environment for viral infection and integration, i.e.: addition of supplementary growth factors. Altering the time of viral transfection throughout the differentiation process may also play a significant role in affecting the efficiency of altering endogenous gene expression. For example, inducing HNF4 α knock down at day 6 may prove to be more successful, as HNF4 α expression can be reduced before peak levels are reached within the cells. The phenotype, in turn, may be more defined.

Nevertheless, another method could be implemented to overcome these issues, the generation of stable cell lines. Generating stable cell lines expressing the required transcripts involves the initial transfection of the cells using the lentiviral vectors as previously described, followed by specific antibiotic selection. This ensures that every surviving cell contains the integrated form of the LV in addition to the stable and constitutive expression of the gene or sequence of interest as well as the resistance gene required for survival. The stable cell lines generated could be inducible, resulting in the ability to initiate or suppress gene expression at any point in time. The mechanism relies on the continuous expression of the Tet repressor in the host cell, which binds as a homo-dimer to the promoter region of the lentivector, previously integrated into the genome, inhibiting its expression. The addition of tetracycline results in the high affinity binding to the Tet repressor initiating a conformational change. Consequently, the Tet repressor is no longer able to bind to

the $TetO_2$ sequence on the viral promoter, inducing the expression of the required transcript. (Yao et al., 1998) The creation of a more stable inducible gene manipulation system will result in an ideal tool for manipulating the differentiation process altering the end production. This technique is translatable to other living system such as primary human hepatocytes, which in turn could be stabilized for long-term survival in culture, a property necessary for the use in therapeutic and research applications.

Generating an indefinite supply of functional hepatocytes provides an indispensible tool essential for industrial and clinical uses such as drug discovery and disease modelling. The provision of clinically relevant cell types will in turn contribute to our understanding of developmental mechanisms resulting in use of these cells for cellular therapy (Alison et al., 2007, Asahina et al., 2006). Functional hepatocytes are a scarce resource, with the majority of sources suffering from limited viability in long-term culture.

A large number of factors have been attributed to efficiently differentiating and maintaining hepatocytes in culture (Agarwal et al., 2008, Hannoun et al., 2010, Wege et al., 2003); however, this thesis isolated HNF4 α as a vital factor in regulating hepatic viability. Interestingly, studies have shown that the levels of HNF4 α are significantly reduced in human and rodent hepatocellular carcinomas, which results in the increase in cellular proliferation and dedifferentiation, in addition to tumour metastasis (Lazarevich et al., 2010). Therefore, HNF4 α is not only required for healthy development and adult liver function, but is also a vital factor in disease such

as renal (Lucas et al., 2005) and liver cancers (Lazarevich et al., 2004) and pancreatic disorders such as MODY1 (Ellard S FAU - Colclough et al.,).

HNF4 α stability was directly associated with poly-SUMOylation, which may have resulted in RNF4 mediated ubiquitination and subsequent degradation via the 26S proteasome. An in depth analysis of the SUMOylation and HNF4 α was attempted within differentiating hESCs. However, due to limitations within the transfection system no defined conclusions could be drawn. Thus, optimization of the transfection technique within hESCs is required to accurately investigate the mechanisms behind HNF4 α and SUMOylation within living systems. This in turn, may provide significant insight required to generate an infinite supply of hESC-derived hepatocytes required for a large number of research, industrial and clinical applications.

CHAPTER SEVEN

DISCUSSION

CHAPTER 7

DISCUSSION

The research presented in this thesis has laid the foundations for detailed studies involved in the development of an in vitro system for the derivation of hepatocytes, whereby the role of HNF4 α was investigated within hESC differentiation into hepatic endoderm.

Human embryonic stem cells (hESCs) are a useful tool for investigating human development and cancer. hESCs are able to self renew indefinitely whilst retaining pluripotentency, (Thomson et al., 1998), as a result they can differentiate into cell types from all three germ layers. hESCs are isolated from the inner cell mass (Thomson et al., 1998), which is responsible for forming the embryo post fertilization. For this reason, in conjunction with their ability to be directly differentiated, hESCs provide good models for studying development. hESCs could provide an indefinite supply of somatic cell types for use in drug discovery, disease modelling, extra-corporeal device construction and eventually cell based therapies (Anna et al., 2005, Cai et al., 2006, Henrik, 2005).

Since the derivation of hESCs from blastocysts in 1998 (Thomson et al., 1998), a large amount of resources have been employed to elucidate the ideal culture conditions for maintaining hESCs in an undifferentiated state whilst retaining pluripotency. Culture techniques range from the use of mouse embryonic fibroblasts (Cai et al., 2006); to provide matrix support and the required nutrients for growth, to

feeder free cultures grown on vitronectin (Braam et al., 2008) or Matrigel™ (Amit et al., 2003) to xeno-free culture systems (Barbara S Mallon et al., 2006). However, most culture conditions suffer from xeno-contamination, lack of definition, limited reproducibility and scalability. Therefore, defining the optimal environmental conditions may allow a better understanding of hESC biology. In an effort to overcome these issues we investigated the potential of a serum free commercial available media, mTeSR® (MT, Stem Cell Technologies, UK), as a replacement for conditioned media (CM), the current gold standard for hESC culture (Hannoun et al., 2010). Our results indicate that MT is a good substitute for culturing hESCs as the cells maintained all embryonic stem cell characteristics. In addition, we observed an improvement in HE generation and function. Therefore, MT standardizes the culture of hESCs providing a suitable platform for efficiently generating functional hepatocytes. Unfortunately, due to limitations associated with batch to batch variation and the requirement for reproducible results, CM was utilized for the remaining investigations.

Functional hepatocytes are a scarce resource with limited viability in long-term culture (Kim et al., 2007). hESC derived HE may in turn alleviate the requirement for the utilization of primary hepatocytes in a number of clinical and research applications (Agarwal et al., 2008, Baharvand et al., 2008). Prior to replacing PHH, it is essential that hESC derived HE exhibit the required hepatic characteristics comparable to freshly isolated PHH. In depth transcriptional analysis of differentiating hESCs using SOLEXA, which verified a gene expression pattern analogous to the hepatic gene expression observed in the developing liver, (Zaret,

2000, Zaret, 2001, Cereghini, 1996, Costa et al., 2003). Functional and transcriptional analysis reinforces the employment of hESC derived HE for a variety of applications. However, further research is vital for refining the differentiation model, resulting in improved hepatocyte function and viability. Hepatic endoderm viability is the most significant limitation with use in all sources of HE. Primary human hepatocytes and hESC derived HE suffer from de-differentiation and death when maintained in long-term culture (Kim et al., 2007). This observation coincides with a decrease in levels of HNF4 α, indicative of a possible role in regulating HE stability. Knock down of HNF4 α during development results in embryo lethality (Chen et al., 1994) and metabolic disorders within the adult liver. Studies have also shown that over-expression of HNF4 α can drive transdifferentiation from hematopoietic cells into hepatocytes (Khurana et al., 2010) and that bone marrow cells relocate to the liver in response to injury and differentiate into hepatocytes defined by the up regulation of HNF4 α (Terai et al., 2003). The combination of the data indicates the importance of HNF4 α in liver development, differentiation and maintenance (Duncan, 2003, Zhao et al., 2005) (Nagaoka et al., 2010).

In an attempt to improve our differentiation model we studied HNF4 α, a key transcription factor required for maintaining HE function and viability. HNF4 α is a constitutively active highly conserved orphan receptor (Wisely et al., 2002), which has been demonstrated to be an essential transcription factor required for healthy liver development and adult liver metabolism (Chen et al., 1994, Lemaigre et al., 2004). As such, we investigated the expression pattern of HNF4 α within our differentiation system. We observed a gradual decrease in HNF4 α expression as

hepatic differentiation proceeded. In the developing liver, levels of HNF4 α decrease in order to stimulate the expression of mature factors; however, post natal, HNF4 α expression is stabilized to preserve normal liver function (Cereghini, 1996, Duncan, 2003, Lemaigre et al., 2004, Runge et al., 1998).

Post translational modifications play important roles in various cellular processes. HNF4 α function is regulated by phosphorylation and Acetylation. However, another important PTM known to alter the function of a variety of proteins is SUMOylation. SUMOylation is an important regulator in a number of cellular processes such as transcription, cell cycle regulation, degradation and differentiation (Johnson, 2004, Kroetz, 2005). To date, the mechanism behind the regulation of HNF4 α stability and function via SUMOylation remains unknown. Therefore, we investigated the relationship between SUMO and HNF4 α within our differentiation model.

HNF4 α was SUMO modified in vitro on lysine residue 365, conserved between all six variants, whereby the decrease in HNF4 α expression throughout differentiation coincided with an increase in the levels of SUMO 2 modification, indicative of poly SUMOylation. Interestingly, Tatham and colleagues established that the poly-SUMOylation of the promyelocytic leukaemia protein (PML) resulted in its ubiquitination, mediated by RNF4, before its degradation via the 26S proteasome (Tatham et al., 2008). In our in vitro model, RNF4 was expressed as the levels of HNF4 α decreased following poly-SUMOylation. This process could be inhibited using a proteasome 26S inhibitor, MG132.

In an attempt to dissect the biological mechanism, we employed the lentivirus technology. The LVs were packaged and used to transfect C3A cells, a hepatoma cell line. Once the technology was optimized, the LV constructs were used to infect hESCs at different time points during cellular differentiation. Despite the effective vectors, no defined conclusions could be drawn. Therefore, further refinement of the technology is required before elucidating the mechanisms within differentiating hESCs.

One method to overcome this issue is the creation of stable cell lines, which could also be inducible. This technique involves antibiotic selection, whereby the remaining surviving cells express the integrated lentiviral vector containing the specific transgene, in addition to the resistance gene, as it is required for survival. Continuous selection ensures that each cell within the population expresses the LV. Transgene expression can also be regulated by the Tet repressor creating inducible stable cell lines. Inducible systems function by inhibiting the expression of the specific transgene by binding the Tet repressor, which is constitutively expressed, onto the lentivirus promoter/enhancer region. The addition of tetracycline results in the binding and conformational change of the Tet repressor, thus promoting the expression of the required transgene. This mechanism in turn can be used to regulate the expression or repression of the gene of interest. Therefore, employing the above technique would be the next step in this investigation. The resulting hESC population will express the required transgene and therefore defined conclusions could be drawn upon manipulation of gene expression. In addition, due to the inducible nature of the system, optimal time points for induction may be developed, further refining the

differentiation system. Optimising hESC viral transfection and successfully manipulating gene expression patterns will, in turn, provide an effective platform that may be translated into various cellular models and differentiation systems, vital in research and clinical applications.

Hepatocyte differentiation is a complex process requiring regulation at multiple levels. This thesis highlights the importance of SUMOylation in creating a stable model for human liver biology. The next stage of this project is to generate inducible stable cells lines which will allow function dissection of the relationship between SUMOylation and HNF4 α within differentiating hESCs. This mechanism in turn may provide insight into techniques that can be used to refine the differentiation model thus providing an indefinite supply of functional viable hepatic endoderm from both hESCs and induced pluripotent stem cells (iPSCs). iPSCs are generated from terminally differentiated somatic cells, whereby pluripotency factors such as Oct4 and Nanog are over expressed, resulting in the reversion of the phenotype into an embryonic stem cell like state (Takahashi et al., 2007). Utilization of iPSCs alleviates the controversy associated with hESCs as they by-pass the requirement of the embryo (Chakraborty et al., 2010). In addition, potential use of iPSCs in therapy reduces the complications attributed to immuno-rejection (Chakraborty et al., 2010). Therefore, iPSCs provide a good platform for modelling hepatic differentiation resulting in potential use in a variety of clinical and research applications. The resulting HE will initially be utilized for disease modelling and drug discovery. Further advances in the technology will afford the opportunity for use in extra corporeal devices with the inevitable employment in cellular therapy once the system

is completely refined. Advancements in each of these fields have been documented; some examples have been described below.

A significant number of differentiation protocols have been successfully developed, resulting in the generation of functional hepatocytes. However, most of the techniques suffer from limitations, specifically complications preventing their use in cell-based therapies. These include xeno-contamination, lack of definition and an impure differentiated population. However, a novel strategy has been established that has efficiently derived hepatocytes directly from mouse fibroblasts using defined factors, bypassing the requirement of embryonic stem cells and iPSCs (Huang et al., 2011). Mouse tail-tip fibroblasts were transduced using lentiviral vectors expressing Gata4, HNF1 α and Fox3a and inactivated $p19^{Arf}$ (Huang et al., 2011). The resulting hepatocytes displayed the correct morphology in conjugation with hepatic specific gene expression and metabolic functions (Huang et al., 2011). The investigation contributed a novel strategy for generating functional hepatocytes directly from terminally differentiated cells, ideal for sourcing an indefinite supply of HE. This technique may be translated into various model systems, human, advancing the utilization of hepatocytes in clinical and therapeutic applications.

Reproducible and scalable models are vital for applications such as drug discovery as it may alleviate the high costs associated with animal tests and clinical trials. An accurate developmental model generated from hESCs derived HE will aid in treating liver disease, which is translatable into other model systems (Sokal, 2011). iPSCs have also been utilized for disease modelling, whereby iPSCs are formed from

patients suffering from various liver disorders (Rashid et al., 2010). Once differentiated, the hepatocytes display phenotypic characteristics of each disease, such as misfolded α1-antitrypsin in the endoplasmic reticulum demonstrated in patients suffering from α1-antitrypsin deficiency and reduced LDL receptor mediated cholesterol uptake from patients suffering from familial hypercholesterolemia (Rashid et al., 2010). In addition, generating HE from induced pluripotent stem cells (iPSCs) depicting individual polymorphic variations within metabolic genes will enhance our understanding of drug metabolism and will improve drug development and toxicology analysis within pharmaceutical companies (Rashid et al., 2010).

Another potential use of iPSC and hESC generated HE is in extra corporeal devices. Extra corporeal devices, the bioartificial liver, have been developed as a support mechanism in acute liver failure cases relieving stress from the damaged liver (Wang et al., 2010). The requirement for a constant supply of functional hepatocytes is vital for such a treatment. Furthermore, provision of functional hepatic endoderm may result in the identification of novel biomarkers and potential drug targets, which will provide significant advancements in both the research and clinical fields (Sokal, 2011).

Formation of iPSCs from individual patients with subsequent differentiation into HE and hESC derived HE, will provide a good tool for cell based therapies. However, the employment of iPSC generated HE reduces implications associated with immuno-rejection and should result in improved cellular engraftment. Liu and colleagues have demonstrated a proof of principle for the use of differentiated HE in

cell therapy. The results demonstrated successful integration of iPSC lines, generated from each germ layer, within the cirrhotic liver of mice, resulting in tissue repopulation and efficient hepatic function, indicated by the presence of human serum proteins within the mouse blood (Liu et al., 2011). The investigation confirms the possibility of the technique as a possible treatment for liver disease in the future in addition to demonstrating that iPSCs generated from every origin can successfully be differentiated into hepatocytes, despite varying methylation patterns defining their epigenetic memory (Liu et al., 2011).

Refining hESCs and iPSCs hepatic endoderm differentiation is vital for its subsequent use in industry, clinic and therapy. This thesis has provided novel evidence of the expression and regulation mechanism of HNF4 α, an essential hepatic transcription factor, within human hepatocyte differentiation. Further investigation into the mechanism and the subsequent effects of HNF4 α regulation in hepatic differentiation may provide insight for the improvement of the differentiation system, translatable into other cellular models.

BIBLIOGRAPHY

Agarwal S, Holton KL, Lanza R (2008) Efficient Differentiation of Functional Hepatocytes From Human Embryonic Stem Cells. *Stem Cells* **26** (5): 1117 - 1127

Alison MR, Choong C, Lim S (2007) Application of Liver Stem Cells for Cell Therapy. *Semin Cell Dev Biol* **18** (6): 819 - 826

Allegrucci C, Young LE (2007) Differences Between Human Embryonic Stem Cell Lines. *Human Reproduction Update* **13** (2): 103 - 120

Alnemri ES, Livingston DJ, Nicholson DW, Salvesen G, Thornberry NA, Wong WW, Yuan J (1996) Human ICE/CED-3 Protease Nomenclature. *Cell* **87** (2): 171 -

Amit M, Itskovitz-Eldor J (2006) Feeder-Free Culture of Human Embryonic Stem Cells. *Methods Enzymol* **420** 37 - 49

Amit M, Margulets V, Segev H, Shariki K, Laevsky I, Coleman R, Itskovitz-Eldor J (2003) Human Feeder Layers for Human Embryonic Stem Cells. *Biol Reprod* **68** (6): 2150 - 2156

Amit M, Shariki C, Margulets V, Itskovitz-Eldor J (2004) Feeder Layer- and Serum-Free Culture of Human Embryonic Stem Cells. *Biol Reprod* **70** (3): 837 - 845

Anna MW, Kenneth RB (2005) Embryonic Stem Cells: Prospects for Developmental Biology and Cell Therapy. *Physiological Review* **85** 635 - 678

Asahina K, Teramoto K, Teraoka H (2006) Embryonic Stem Cells: Hepatic Differentiation and Regenerative Medicine for the Treatment of Liver Disease. *Curr Stem Cell Res Ther* **1** (2): 139 - 156

Baharvand H, Hashemi SM, Kazemi AS, Farrokhi A (2006) Differentiation of Human Embryonic Stem Cells into Hepatocytes in 2D and 3D Culture Systems in Vitro. *Int J Dev Biol* **50** (7): 645 - 652

Baharvand H, Hashemi SM, Shahsavani M (2008) Differentiation of Human Embryonic Stem Cells into Functional Hepatocyte-Like Cells in a Serum-Free Adherent Culture Condition. *Differentiation* **76** (5): 465 - 477

Barbara S Mallon, Kye YP, Kevin GC, Rebecca S Hamilton, Ronald DM (2006) Toward Xeno Free Culture in Human Embryonic Stem Cells. *IJBCB* (38): 1063 - 1075

Basma H, Soto-Gutierrez A, Yannam GR, Liu L, Ito R, Yamamoto T, Ellis E, Carson SD, Sato S, Chen Y, Muirhead D, Navarro-Alvarez N, Wong RJ, Roy-Chowdhury J, Platt JL, Mercer DF, Miller JD, Strom SC, Kobayashi N, Fox

IJ (2009) Differentiation and Transplantation of Human Embryonic Stem Cell-Derived Hepatocytes. *Gastroenterology* **136** (3): 990 - 999

Beattie GM, Lopez AD, Bucay N, Hinton A, Firpo MT, King CC, Hayek A (2005) Activin A Maintains Pluripotency of Human Embryonic Stem Cells in the Absence of Feeder Layers. *Stem Cells* **23** (4): 489 - 495

Behnia K, Bhatia S, Jastromb N, Balis U, Sullivan S, Yarmush M, Toner M (2000) Xenobiotic Metabolism by Cultured Primary Porcine Hepatocytes. *Tissue Eng* **6** (5): 467 - 479

Bezerra JA (1998) Liver Development: a Paradigm for Hepatobiliary Disease in Later Life. *Semin Liver Dis* **18** (3): 203 - 216

Bogan AA, Dallas-Yang Q, Ruse MD, Maeda Y, Jiang G, Nepomuceno L, Scanlan TS, Cohen FE, Sladek FM (2000) Analysis of Protein Dimerization and Ligand Binding of Orphan Receptor HNF4[Alpha]. *Journal of molecular biology* **302** (4): 831 - 851

Bonavita AG, Quaresma K, Cotta-de-Almeida V, Pinto MA, Saraiva RM, Alves LA (2010) Hepatocyte Xenotransplantation for Treating Liver Disease. *Xenotransplantation.* **17** (3): 181 - 187

Bossard P, Zaret KS (1998) GATA Transcription Factors As Potentiators of Gut Endoderm Differentiation. *Development (Cambridge, England)* **125** (24): 4909 - 4917

Braam SR, Zeinstra L, Litjens S, Ward-van Oostwaard D, van den BS, van Laake L, Lebrin F, Kats P, Hochstenbach R, Passier R, Sonnenberg A, Mummery CL (2008) Recombinant Vitronectin Is a Functionally Defined Substrate That Supports Human Embryonic Stem Cell Self-Renewal Via Alphavbeta5 Integrin. *Stem Cells* **26** (9): 2257 - 2265

Briancon N, Bailly A, Clotman F, Jacquemin P, Lemaigre FP, Weiss MC (2004) Expression of the Alpha7 Isoform of Hepatocyte Nuclear Factor (HNF) 4 Is Activated by HNF6/OC-2 and HNF1 and Repressed by HNF4alpha1 in the Liver. *J Biol Chem* **279** (32): 33398 - 33408

Briancon N, Weiss MC (2006) In Vivo Role of the HNF4alpha AF-1 Activation Domain Revealed by Exon Swapping. *EMBO J* **25** (6): 1253 - 1262

Brimble SN, Zeng X, Weiler DA, Luo Y, Liu Y, Lyons IG, Freed WJ, Robins AJ, Rao MS, Schulz TC (2004) Karyotypic Stability, Genotyping, Differentiation, Feeder-Free Maintenance, and Gene Expression Sampling in Three Human Embryonic Stem Cell Lines Derived Prior to August 9, 2001. *Stem Cells Dev* **13** (6): 585 - 597

Brink TC, Sudheer S, Janke D, Jagodzinska J, Jung M, Adjaye J (2008) The Origins of Human Embryonic Stem Cells: a Biological Conundrum. *Cells Tissues Organs* **188** (1-2): 9 - 22

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L (2007) Derivation of Pluripotent Epiblast Stem Cells From Mammalian Embryos. *Nature* **448** (7150): 191 - 195

Buecker C, Chen HH, Polo JM, Daheron L, Bu L, Barakat TS, Okwieka P, Porter A, Gribnau J, Hochedlinger K, Geijsen N (2010) A Murine ESC-Like State Facilitates Transgenesis and Homologous Recombination in Human Pluripotent Stem Cells. *Cell stem cell* **6** (6): 535 - 546

Burke ZrD, Thowfeequ S, Tosh D (2006) Liver Specification: A New Role for Wnts in Liver Development. **16** (17): R688 - R690

Burnett G, Kennedy EP (1954) The Enzymatic Phosphorylation of Proteins. *The Journal of biological chemistry* **211** (2): 969 - 980

Burns JC, Friedmann T, Driever W, Burrascano M, Yee JK (1993) Vesicular Stomatitis Virus G Glycoprotein Pseudotyped Retroviral Vectors: Concentration to Very High Titer and Efficient Gene Transfer into Mammalian and Nonmammalian Cells. *Proc Natl Acad Sci U S A* **90** (17): 8033 - 8037

Butler D (1998) Last Chance to Stop and Think on Risks of Xenotransplants. *Nature* **391** (6665): 320 - 324

Cai J, Zhao Y, Liu Y, Ye F, Song Z, Qin H, Meng S, Chen Y, Zhou R, Song X, Guo Y, Ding M, Deng H (2007) Directed Differentiation of Human Embryonic Stem Cells into Functional Hepatic Cells. *Hepatology* **45** (5): 1229 - 1239

Cai J, Chen J, Liu Y, Miura T, Luo Y, Loring JF, Freed WJ, Rao MS, Zeng X (2006) Assessing Self-Renewal and Differentiation in Human Embryonic Stem Cell Lines. *Stem cells (Dayton, Ohio)* **24** (3): 516 - 530

Calmont A, Wandzioch E, Tremblay KD, Minowada G, Kaestner KH, Martin GR, Zaret KS (2006) An FGF Response Pathway That Mediates Hepatic Gene Induction in Embryonic Endoderm Cells. *Developmental cell* **11** (3): 339 - 348

Campos PB, Sartore RC, Abdalla SN, Rehen SK (2009) Chromosomal Spread Preparation of Human Embryonic Stem Cells for Karyotyping. *J Vis Exp* (31):

Cereghini S (1996) Liver-Enriched Transcription Factors and Hepatocyte Differentiation. *FASEB J* **10** (2): 267 - 282

Chakraborty C, Shah KD, Cao WG, Hsu CH, Wen ZH, Lin CS (2010) Potentialities of Induced Pluripotent Stem (IPS) Cells for Treatment of Diseases. *Curr Mol Med* **10** (8): 756 - 762

Charron F, Tsimiklis G, Arcand M, Robitaille L, Liang Q, Molkentin JD, Meloche S, Nemer M (2001) Tissue-Specific GATA Factors Are Transcriptional Effectors of the Small GTPase RhoA. *Genes Dev* **15** (20): 2702 - 2719

Chen G, Gulbranson DR, Hou Z, Bolin JM, Ruotti V, Probasco MD, Smuga-Otto K, Howden SE, Diol NR, Propson NE, Wagner R, Lee GO, Antosiewicz-Bourget J, Teng JM, Thomson JA (2011) Chemically defined conditions for human iPSC derivation and culture. *Nature Methods* **8** (5): 424-429

Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, Darnell JE (1994) Disruption of the HNF-4 Gene, Expressed in Visceral Endoderm, Leads to Cell Death in Embryonic Ectoderm and Impaired Gastrulation of Mouse Embryos. *Genes & development* **8** (20): 2466 - 2477

Cheng L, Hammond H, Ye Z, Zhan X, Dravid G (2003) Human Adult Marrow Cells Support Prolonged Expansion of Human Embryonic Stem Cells in Culture. *Stem Cells* **21** (2): 131 - 142

Chiba H, Itoh T, Satohisa S, Sakai N, Noguchi H, Osanai M, Kojima T, Sawada N (2005) Activation of P21CIP1/WAF1 Gene Expression and Inhibition of Cell Proliferation by Overexpression of Hepatocyte Nuclear Factor-4alpha. *Exp Cell Res* **302** (1): 11 - 21

Chin AC, Fong WJ, Goh LT, Philp R, Oh SK, Choo AB (2007) Identification of Proteins From Feeder Conditioned Medium That Support Human Embryonic Stem Cells. *J Biotechnol* **130** (3): 320 - 328

Chin AC, Padmanabhan J, Oh SK, Choo AB (2010) Defined and Serum-Free Media Support Undifferentiated Human Embryonic Stem Cell Growth. *Stem Cells Dev* **19** (6): 753 - 761

Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS (2002) Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GA4. *Molecular cell* **9** (2): 279 - 289

Clotman F, Jacquemin P, Plumb-Rudewiez N, Pierreux CE, Van der SP, Dietz HC, Courtoy PJ, Rousseau GG, Lemaigre FP (2005) Control of Liver Cell Fate Decision by a Gradient of TGF Beta Signalling Modulated by Onecut Transcription Factors. *Genes Dev* **19** (16): 1849 - 1854

Coffinier C, Gresh L, Fiette L, Tronche Fo, Sch€ªtz G, Babinet C, Pontoglio M, Yaniv M, Barra J (2002) Bile System Morphogenesis Defects and Liver Dysfunction Upon Targeted Deletion of HNF1beta. *Development (Cambridge, England)* **129** (8): 1829 - 1838

Costa RH, Kalinichenko VV, Holterman AX, Wang X (2003) Transcription Factors in Liver Development, Differentiation, and Regeneration. *Hepatology* **38** (6): 1331 - 1347

Crook JM, Peura T, Kravets L, Bosman AG, Buzzard JJ, Horne R (2007) The Generation of Six Clinical-Grade Human Embryonic Stem Cell Lines. *Cell stem cell* **1** (5): 490 - 494

Cuddapah S, Barski A, Cui K, Schones DE, Wang Z, Wei G, Zhao K (2009) Native Chromatin Preparation and Illumina/Solexa Library Construction. *Cold Spring Harb.Protoc* **2009** (6): db -

Dame C, Sola MC, Lim KC, Leach KM, Fandrey J, Ma Y, Knopfle G, Engel JD, Bungert J (2004) Hepatic Erythropoietin Gene Regulation by GATA-4. *J Biol Chem* **279** (4): 2955 - 2961

Decker CJ, Parker R (1994) Mechanisms of MRNA Degradation in Eukaryotes. *Trends Biochem Sci* **19** (8): 336 - 340

Desterro JM, Rodriguez MS, Hay RT (1998) SUMO-1 Modification of IkappaBalpha Inhibits NF-KappaB Activation. *Molecular cell* **2** (2): 233 - 239

Desterro JM, Rodriguez MS, Kemp GD, Hay RT (1999) Identification of the Enzyme Required for Activation of the Small Ubiquitin-Like Protein SUMO-1. *J Biol Chem* **274** (15): 10618 - 10624

Desterro JM, Thomson J, Hay RT (1997) Ubch9 Conjugates SUMO but Not Ubiquitin. *FEBS letters* **417** (3): 297 - 300

Dhe-Paganon S, Duda K, Iwamoto M, Chi YI, Shoelson SE (2002) Crystal Structure of the HNF4x¤ Ligand Binding Domain in Complex With Endogenous Fatty Acid Ligand. *Journal of Biological Chemistry* **277** (41): 37973 - 37976

Dou H, Huang C, Van Nguyen T, Lu LS, Yeh ETH (2011) SUMOylation and De-SUMOylation in Response to DNA Damage. *FEBS letters* **In Press, Corrected Proof**

Drewes T, Senkel S, Holewa B, Ryffel GU (1996) Human Hepatocyte Nuclear Factor 4 Isoforms Are Encoded by Distinct and Differentially Expressed Genes. *Mol Cell Biol* **16** (3): 925 - 931

Du Y, Han R, Wen F, Ng San San S, Xia L, Wohland T, Leo HL, Yu H (2008) Synthetic Sandwich Culture of 3D Hepatocyte Monolayer. *Biomaterials* **29** (3): 290 - 301

Duda K, Chi YI, Shoelson SE (2004) Structural Basis for HNF-4alpha Activation by Ligand and Coactivator Binding. *J Biol Chem* **279** (22): 23311 - 23316

Duncan SA (2003) Mechanisms Controlling Early Development of the Liver. *Mechanisms of development* **120** (1): 19 - 33

Eeckhoute J, Moerman E, Bouckenooghe T, Lukoviak B, Pattou F, Formstecher P, Kerr-Conte J, Vandewalle B, Laine B (2003) Hepatocyte Nuclear Factor

4{Alpha} Isoforms Originated From the P1 Promoter Are Expressed in Human Pancreatic {Beta}-Cells and Exhibit Stronger Transcriptional Potentials Than P2 Promoter-Driven Isoforms. *Endocrinology* **144** (5): 1686 - 1694

Ellard S FAU - Colclough K , Colclough K Mutations in the Genes Encoding the Transcription Factors Hepatocyte Nuclear Factor 1 Alpha (HNF1A) and 4 Alpha (HNF4A) in Maturity-Onset Diabetes of the Young. (1098-1004 (Electronic)):

Fiegel HC, Kaufmann PM, Bruns H, Kluth D, Horch RE, Vacanti JP, Kneser U (2008) Hepatic Tissue Engineering: From Transplantation to Customized Cell-Based Liver Directed Therapies From the Laboratory. *J Cell Mol Med* **12** (1): 56 - 66

Fiegel HC, Lange C, Kneser U, Lambrecht W, Zander AR, Rogiers X, Kluth D (2006) Fetal and Adult Liver Stem Cells for Liver Regeneration and Tissue Engineering. *Journal of cellular and molecular medicine* **10** (3): 577 - 587

Filippi C, Keatch SA, Rangar D, Nelson LJ, Hayes PC, Plevris JN (2004) Improvement of C3A Cell Metabolism for Usage in Bioartificial Liver Support Systems. *J Hepatol.* **41** (4): 599 - 605

Finer JJ, McMullen MD (1990) Transformation of Cotton (≪i≫Gossypium Hirsutum≪/i≫ L.) Via Particle Bombardment. *Plant Cell Reports* **8** (10): 586 - 589

Fletcher J, Cui W, Samuel K, Black JR, Hannoun Z, Currie IS, Terrace JD, Payne C, Filippi C, Newsome P, Forbes SJ, Ross JA, Iredale JP, Hay DC (2008) The Inhibitory Role of Stromal Cell Mesenchyme on Human Embryonic Stem Cell Hepatocyte Differentiation Is Overcome by Wnt3a Treatment. *Cloning Stem Cells* **10** (3): 331 - 339

Frazier AE, Kiu C, Stojanovski D, Hoogenraad NJ, Ryan MT (2006) Mitochondrial Morphology and Distribution in Mammalian Cells. *Biological chemistry* **387** (12): 1551 - 1558

Frost J, Monk D, Moschidou D, Guillot PV, Stanier P, Minger SL, Fisk NM, Moore HD, Moore GE (2011) The Effects of Culture on Genomic Imprinting Profiles in Human Embryonic and Fetal Mesenchymal Stem Cells. *Epigenetics.* **6** (1): 52 - 62

Glozak MA, Sengupta N, Zhang X, Seto E (2005) Acetylation and Deacetylation of Non-Histone Proteins. *Gene* **363** 15 - 23

Gonzalez FJ (2008) Regulation of Hepatocyte Nuclear Factor 4 Alpha-Mediated Transcription. *Drug Metab Pharmacokinet* (1880-0920 (Electronic)):

Grewal SIS, Rice JC (2004) Regulation of Heterochromatin by Histone Methylation and Small RNAs. *Current opinion in cell biology* **16** (3): 230 - 238

Guo H, Gao C, Mi Z, Wai PY, Kuo PC (2006) Phosphorylation of Ser158 Regulates Inflammatory Redox-Dependent Hepatocyte Nuclear Factor-4Iñ Transcriptional Activity. *Biochemical Journal* **394** (2): 379 - 387

Gupta RK, Gao N, Gorski RK, White P, Hardy OT, Rafiq K, Brestelli JE, Chen G, Stoeckert CJ, Jr., Kaestner KH (2007) Expansion of Adult Beta-Cell Mass in Response to Increased Metabolic Demand Is Dependent on HNF-4alpha. *Genes Dev* **21** (7): 756 - 769

Hadzopoulou-Cladaras M, Kistanova E, Evagelopoulou C, Zeng S, Cladaras C, Ladias JA (1997) Functional Domains of the Nuclear Receptor Hepatocyte Nuclear Factor 4. *J Biol Chem* **272** (1): 539 - 550

Han MK, Song EK, Guo Y, Ou X, Mantel C, Broxmeyer HE (2008) SIRT1 Regulates Apoptosis and Nanog Expression in Mouse Embryonic Stem Cells by Controlling P53 Subcellular Localization. *Cell stem cell* **2** (3): 241 - 251

Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, Cassady JP, Muffat J, Carey BW, Jaenisch R (2010) Human Embryonic Stem Cells With Biological and Epigenetic Characteristics Similar to Those of Mouse ESCs. *Proceedings of the National Academy of Sciences* **107** (20): 9222 - 9227

Hannoun Z, Filippi C, Sullivan G, Hay DC, Iredale JP (2010) Hepatic Endoderm Differentiation From Human Embryonic Stem Cells. *Curr.Stem Cell Res.Ther.* **5** (3): 233 - 244

Hannoun Z, Fletcher J, Greenhough S, Medine C, Samuel K, Sharma R, Pryde A, Black JR, Ross JA, Wilmut I, Iredale JP, Hay DC (2010) The Comparison Between Conditioned Media and Serum-Free Media in Human Embryonic Stem Cell Culture and Differentiation. *Cellular Reprogramming* **12** (2 DOI - 10.1089/cell.2009.0099): 133 - 140

Hannoun Z, Greenhough S, Jaffray E, Hay RT, Hay DC (2010) Post-Translational Modification by SUMO. *Toxicology* **278** (3): 288 - 293

Harder Z, Zunino R, McBride H (2004) Sumo1 Conjugates Mitochondrial Substrates and Participates in Mitochondrial Fission. *Current biology : CB* **14** (4): 340 - 345

Harish S, Khanam T, Mani S, Rangarajan P (2001) Transcriptional Activation by Hepatocyte Nuclear Factor-4 in a Cell-Free System Derived From Rat Liver Nuclei. *Nucleic Acids Res* **29** (5): 1047 - 1053

Harries LW, Brown JE, Gloyn AL (2009) Species-Specific Differences in the Expression of the HNF1A, HNF1B and HNF4A Genes. *PLoS ONE* **4** (11): e7855 -

Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, Steele A, NjÃ¸lstad PlR, Ellard S, Hattersley AT (2008) The Diabetic Phenotype in HNF4A Mutation

Carriers Is Moderated By the Expression of HNF4A Isoforms From the P1 Promoter During Fetal Development. *Diabetes* **57** (6): 1745 - 1752

Hartwell L, Hood L, Goldberg M, Reynolds A, Silver L (2010) Genetics: From Genes to Genomes. *4th*

Hattori H, Imai H, Kirai N, Furuhama K, Sato O, Konishi K, Nakagawa Y (2007) Identification of a Responsible Promoter Region and a Key Transcription Factor, CCAAT/Enhancer-Binding Protein Epsilon, for Up-Regulation of PHGPx in HL60 Cells Stimulated With TNF Alpha. *Biochem J* **408** (2): 277 - 286

Hay DC, Fletcher J, Payne C, Terrace JD, Gallagher RC, Snoeys J, Black JR, Wojtacha D, Samuel K, Hannoun Z, Pryde A, Filippi C, Currie IS, Forbes SJ, Ross JA, Newsome PN, Iredale JP (2008) Highly Efficient Differentiation of HESCs to Functional Hepatic Endoderm Requires ActivinA and Wnt3a Signalling. *Proc.Natl.Acad.Sci.U.S.A* **105** (34): 12301 - 12306

Hay DC, Zhao D, Fletcher J, Hewitt ZA, McLean D, Urruticoechea-Uriguen A, Black JR, Elcombe C, Ross JA, Wolf R, Cui W (2008) Efficient Differentiation of Hepatocytes From Human Embryonic Stem Cells Exhibiting Markers Recapitulating Liver Development in Vivo. *Stem Cells* **26** (4): 894 - 902

Hay DC, Zhao D, Ross A, Mandalam R, Lebkowski J, Cui W (2007) Direct Differentiation of Human Embryonic Stem Cells to Hepatocyte-Like Cells Exhibiting Functional Activities. *Cloning Stem Cells* **9** (1): 51 - 62

Hay DC, Pernagallo S, Diaz-Mochon JJ, Medine CN, Greenhough S, Hannoun Z, Schrader J, Black JR, Fletcher J, Dalgetty D, Thompson AI, Newsome PN, Forbes SJ, Ross JA, Bradley M, Iredale JP Unbiased Screening of Polymer Libraries to Define Novel Substrates for Functional Hepatocytes With Inducible Drug Metabolism. *Stem Cell Research* **In Press, Corrected Proof**

Hay RT, Vuillard L, Desterro JM, Rodriguez MS (1999) Control of NF-Kappa B Transcriptional Activation by Signal Induced Proteolysis of I Kappa B Alpha. *Philos.Trans R.Soc Lond B Biol Sci* **354** (1389): 1601 - 1609

Hayhurst GP, Strick-Marchand H, Mulet C, Richard AF, Morosan S, Kremsdorf D, Weiss MC (2008) Morphogenetic Competence of HNF4 Alpha-Deficient Mouse Hepatic Cells. *J Hepatol.* **49** (3): 384 - 395

Hayhurst G, Lee YH, Lambert G, Ward J, Gonzalez FJ (2001) Hepatocyte Nuclear Factor 4alpha (Nuclear Receptor 2A1) Is Essential for Maintenance of Hepatic Gene Expression and Lipid Homeostasis. *Molecular and cellular biology* (0270-7306 (Print)):

He X, Ruthann MP, Dong S Li, Xuezhu F, Tenneille L, James AT (2005) Basic FGF and Supression of BMP Signalling Sustain Undiffrentiated Proliferation of Human ES Cells. *Nature Methods* **2** 185 - 190

Henrik S (2005) Human Embryonic Stem Cells: Origin, Properties and Applications. *APMIS* (113): 743 - 750

Hertz R, Magenheim J, Berman I, Bar-Tana J (1998) Fatty Acyl-CoA Thioesters Are Ligands of Hepatic Nuclear Factor-4alpha. *Nature* **392** (6675): 512 - 516

Hoffman LM, Carpenter MK (2005) Characterization and Culture of Human Embryonic Stem Cells. *Nat Biotechnol* **23** (6): 699 - 708

Hoffmann A, Roeder RG (1991) Purification of His-Tagged Proteins in Non-Denaturing Conditions Suggests a Convenient Method for Protein Interaction Studies. *Nucleic acids research* **19** (22): 6337 - 6338

Hong Y, Rogers R, Matunis MJ, Mayhew CN, Goodson ML, Park-Sarge OK, Sarge KD, Goodson M (2001) Regulation of Heat Shock Transcription Factor 1 by Stress-Induced SUMO-1 Modification. *The Journal of biological chemistry* **276** (43): 40263 - 40267

Hong YH, Varanasi US, Yang W, Leff T (2003) AMP-Activated Protein Kinase Regulates HNF4Iñ Transcriptional Activity by Inhibiting Dimer Formation and Decreasing Protein Stability. *Journal of Biological Chemistry* **278** (30): 27495 - 27501

Hu XV, Rodrigues TnMA, Tao H, Baker RK, Miraglia L, Orth AP, Lyons GE, Schultz PG, Wu X (2010) Identification of RING Finger Protein 4 (RNF4) As a Modulator of DNA Demethylation Through a Functional Genomics Screen. *Proceedings of the National Academy of Sciences* **107** (34): 15087 - 15092

Huang H, Ruan H, Aw MY, Hussain A, Guo L, Gao C, Qian F, Leung T, Song H, Kimelman D, Wen Z, Peng J (2008) Mypt1-Mediated Spatial Positioning of Bmp2-Producing Cells Is Essential for Liver Organogenesis. *Development (Cambridge, England)* **135** (19): 3209 - 3218

Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L (2011) Induction of Functional Hepatocyte-Like Cells From Mouse Fibroblasts by Defined Factors. *Nature*

Ihara A, Yamagata K, Nammo T, Miura A, Yuan M, Tanaka T, Sladek FM, Matsuzawa Y, Miyagawa J, Shimomura I (2005) Functional Characterization of the HNF4alpha Isoform (HNF4alpha8) Expressed in Pancreatic Beta-Cells. *Biochem Biophys Res Commun* **329** (3): 984 - 990

Ilic D (2006) Culture of Human Embryonic Stem Cells and the Extracellular Matrix Microenvironment. *Regen Med* **1** (1): 95 - 101

Imhof A (2006) Epigenetic Regulators and Histone Modification. *Brief.Funct.Genomic.Proteomic.* **5** (3): 222 - 227

Iyer VV, Yang H, Ierapetritou MG, Roth CM (2010) Effects of Glucose and Insulin on HepG2-C3A Cell Metabolism. *Biotechnol Bioeng* **107** (2): 347 - 356

Jacquemin P, Lemaigre F, Rousseau GG (2003) The Onecut Transcription Factor HNF-6 (OC-1) Is Required for Timely Specification of the Pancreas and Acts Upstream of Pdx-1 in the Specification Cascade. *Developmental biology* **258** (1): 105 - 116

James D, Levine AJ, Besser D, Hemmati-Brivanlou A (2005) TGFbeta/Activin/Nodal Signalling Is Necessary for the Maintenance of Pluripotency in Human Embryonic Stem Cells. *Development* **132** (6): 1273 - 1282

Jean L, Runhua H, Carmen JB, Shih HY, Micheal S (2006) Defined Culture Conditions of Human Embryonic Stem Cells. *PNAS* **103** (15): 5688 - 5693

Jiang G, Lee U, Sladek FM (1997) Proposed Mechanism for the Stabilization of Nuclear Receptor DNA Binding Via Protein Dimerization. *Mol Cell Biol* (0270-7306 (Print)):

Johnson ES (2004) Protein Modification by SUMO. *Annual review of biochemistry* **73** 355 - 382

Josephson R, Sykes G, Liu Y, Ording C, Xu W, Zeng X, Shin S, Loring J, Maitra A, Rao MS, Auerbach JM (2006) A Molecular Scheme for Improved Characterization of Human Embryonic Stem Cell Lines. *BMC Biol* **4** 28 -

Jung J, Zheng M, Goldfarb M, Zaret KS (1999) Initiation of Mammalian Liver Development From Endoderm by Fibroblast Growth Factors. *Science (New York, N.Y.)* **284** (5422): 1998 - 2003

Kaestner KH (2000) The Hepatocyte Nuclear Factor 3 (HNF3 or FOXA) Family in Metabolism. *Trends in endocrinology and metabolism: TEM* **11** (7): 281 - 285

Kagoshima M, Kinoshita T, Matsumoto K, Nakamura T (1992) Developmental Changes in Hepatocyte Growth Factor MRNA and Its Receptor in Rat Liver, Kidney and Lung. *Eur J Biochem* **210** (1): 375 - 380

Kaikkonen S, Jaaskelainen T, Karvonen U, Rytinki MM, Makkonen H, Gioeli D, Paschal BM, Palvimo JJ (2009) SUMO-Specific Protease 1 (SENP1) Reverses the Hormone-Augmented SUMOylation of Androgen Receptor and Modulates Gene Responses in Prostate Cancer Cells. *Molecular Endocrinology* **23** (3): 292 - 307

Kamiya A, Kinoshita T, Miyajima A (2001) Oncostatin M and Hepatocyte Growth Factor Induce Hepatic Maturation Via Distinct Signalling Pathways. *FEBS Lett* **492** (1-2): 90 - 94

Karpova A, Sanna PP, Behnisch T (2006) Involvement of Multiple Phosphatidylinositol 3-Kinase-Dependent Pathways in the Persistence of Late-Phase Long Term Potentiation Expression. *Neuroscience* **137** (3): 833 - 841

Khurana S, Jaiswal AK, Mukhopadhyay A (2010) Hepatocyte Nuclear Factor-4alpha Induces Transdifferentiation of Hematopoietic Cells into Hepatocytes. *J Biol Chem* **285** (7): 4725 - 4731

Kim SH, Akaike T (2007) Epidermal Growth Factor Signalling for Matrix-Dependent Cell Proliferation and Differentiation in Primary Cultured Hepatocytes. *Tissue engineering* **13** (3): 601 - 609

Kinoshita T, Miyajima A (2002) Cytokine Regulation of Liver Development. *Biochim.Biophys Acta* **1592** (3): 303 - 312

Klein TM, Wolf ED, Wu R, Sanford JC (1987) High-Velocity Microprojectiles for Delivering Nucleic Acids into Living Cells. *Nature* **327** (6117): 70 - 73

Kroetz MB (2005) SUMO: a Ubiquitin-Like Protein Modifier. *The Yale journal of biology and medicine* **78** (4): 197 - 201

Kuai XL, Cong XQ, Li XL, Xiao SD (2003) Generation of Hepatocytes From Cultured Mouse Embryonic Stem Cells. *Liver Transpl.* **9** (10): 1094 - 1099

Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A (2007) FGF Stimulation of the Erk1/2 Signalling Cascade Triggers Transition of Pluripotent Embryonic Stem Cells From Self-Renewal to Lineage Commitment. *Development* **134** (16): 2895 - 2902

Kuo CJ, Conley PB, Chen L, Sladek FM, Darnell JE, Jr., Crabtree GR (1992) A Transcriptional Hierarchy Involved in Mammalian Cell-Type Specification. *Nature* **355** (6359): 457 - 461

Kyrmizi IF, Hatzis PF, Katrakili NF, Tronche FF, Gonzalez FJ FAU - Talianidis I, Talianidis I Plasticity and Expanding Complexity of the Hepatic Transcription Factor Network During Liver Development. (0890-9369 (Print)):

Laimins LA, Gruss P, Pozzatti R, Khoury G (1984) Characterization of Enhancer Elements in the Long Terminal Repeat of Moloney Murine Sarcoma Virus. *The Journal of Virology* **49** (1): 183 - 189

Lallemand-Breitenbach V, Jeanne M, Benhenda S, Nasr R, Lei M, Peres L, Zhou J, Zhu J, Raught B, de The H (2008) Arsenic Degrades PML or PML-

RARalpha Through a SUMO-Triggered RNF4/Ubiquitin-Mediated Pathway. *Nat Cell Biol* **10** (5): 547 - 555

Lazarevich NL, Cheremnova OA, Varga EV, Ovchinnikov DA, Kudrjavtseva EI, Morozova OV, Fleishman DI, Engelhardt NV, Duncan SA (2004) Progression of HCC in Mice Is Associated With a Downregulation in the Expression of Hepatocyte Nuclear Factors. *Hepatology* **39** (4): 1038 - 1047

Le Tissier P, Stoye JP, Takeuchi Y, Patience C, Weiss RA (1997) Two Sets of Human-Tropic Pig Retrovirus. *Nature* **389** (6652): 681 - 682

Lee CS, Friedman JR, Fulmer JT, Kaestner KH (2005) The Initiation of Liver Development Is Dependent on Foxa Transcription Factors. *Nature* **435** (7044): 944 - 947

Lee H, Shamy GA, Elkabetz Y, Schofield CM, Harrsion NL, Panagiotakos G, Socci ND, Tabar V, Studer L (2007) Directed Differentiation and Transplantation of Human Embryonic Stem Cell-Derived Motoneurons. *Stem Cells* **25** (8): 1931 - 1939

Lemaigre F, Zaret KS (2004) Liver Development Update: New Embryo Models, Cell Lineage Control, and Morphogenesis. *Current opinion in genetics & development* **14** (5): 582 - 590

Li J, Ning G, Duncan SA (2000) Mammalian Hepatocyte Differentiation Requires the Transcription Factor HNF-4alpha. *Genes Dev* **14** (4): 464 - 474

Lima CD, Reverter D (2008) Structure of the Human SENP7 Catalytic Domain and Poly-SUMO Deconjugation Activities for SENP6 and SENP7. *Journal of Biological Chemistry* **283** (46): 32045 - 32055

Liu H, Kim Y, Sharkis S, Marchionni L, Jang YY (2011) In Vivo Liver Regeneration Potential of Human Induced Pluripotent Stem Cells From Diverse Origins. *Sci Transl.Med* **3** (82): 82ra39 -

Lokmane L, Haumaitre C, Garcia-Villalba P, Anselme I, Schneider-Maunoury S, Cereghini S (2008) Crucial Role of VHNF1 in Vertebrate Hepatic Specification. *Development* **135** (16): 2777 - 2786

Lu HF, Lim WS, Wang J, Tang ZQ, Zhang PC, Leong KW, Chia SM, Yu H, Mao HQ (2003) Galactosylated PVDF Membrane Promotes Hepatocyte Attachment and Functional Maintenance. *Biomaterials* **24** (27): 4893 - 4903

Lu X, Lane DP (1993) Differential Induction of Transcriptionally Active P53 Following UV or Ionizing Radiation: Defects in Chromosome Instability Syndromes? *Cell* **75** (4): 765 - 778

Lucas B, Grigo K, Erdmann S, Lausen J, Klein-Hitpass L, Ryffel GU (2005) HNF4alpha Reduces Proliferation of Kidney Cells and Affects Genes Deregulated in Renal Cell Carcinoma. *Oncogene* **24** (42): 6418 - 6431

Ludwig TE, Bergendahl V, Levenstein ME, Yu J, Probasco MD, Thomson JA (2006) Feeder-Independent Culture of Human Embryonic Stem Cells. *Nat Methods* **3** (8): 637 - 646

Ludwig TE, Levenstein ME, Jones JM, Berggren WT, Mitchen ER, Frane JL, Crandall LJ, Daigh CA, Conard KR, Piekarczyk MS, Llanas RA, Thomson JA (2006) Derivation of Human Embryonic Stem Cells in Defined Conditions. *Nat Biotech* **24** (2): 185 - 187

Macgillivray AJ, Paul J, Threlfall G (1972) Transcriptional Regulation in Eukaryotic Cells. *Adv.Cancer Res* **15** 93 - 162

Mahajan R, Delphin C, Guan T, Gerace L, Melchior F (1997) A Small Ubiquitin-Related Polypeptide Involved in Targeting RanGAP1 to Nuclear Pore Complex Protein RanBP2. *Cell* **88** (1): 97 - 107

Matsumoto K, Miki R, Nakayama M, Tatsumi N, Yokouchi Y (2008) Wnt9a Secreted From the Walls of Hepatic Sinusoids Is Essential for Morphogenesis, Proliferation, and Glycogen Accumulation of Chick Hepatic Epithelium. *Dev Biol* **319** (2): 234 - 247

Mavri-Damelin D, Damelin LH, Eaton S, Rees M, Selden C, Hodgson HJ (2008) Cells for Bioartificial Liver Devices: the Human Hepatoma-Derived Cell Line C3A Produces Urea but Does Not Detoxify Ammonia. *Biotechnol Bioeng* **99** (3): 644 - 651

McLin V, Rankin SA, Zorn AM (2007) Repression of Wnt/Beta-Catenin Signalling in the Anterior Endoderm Is Essential for Liver and Pancreas Development. *Development (Cambridge, England)* **134** (12): 2207 - 2217

Medine CN, Greenhough S, Hay DC (2010) Role of Stem-Cell-Derived Hepatic Endoderm in Human Drug Discovery. *Biochem Soc Trans* **38** (4): 1033 - 1036

Melchior F, Schergaut M, Pichler A (2003) SUMO: Ligases, Isopeptidases and Nuclear Pores. *Trends in biochemical sciences* **28** (11): 612 - 618

Melkoumian Z, Weber JL, Weber DM, Fadeev AG, Zhou Y, Dolley-Sonneville P, Yang J, Qiu L, Priest CA, Shogbon C, Martin AW, Nelson J, West P, Beltzer JP, Pal S, Brandenberger R (2010) Synthetic Peptide-Acrylate Surfaces for Long-Term Self-Renewal and Cardiomyocyte Differentiation of Human Embryonic Stem Cells. *Nat Biotechnol* **28** (6): 606 - 610

Mitalipova MM, Rao RR, Hoyer DM, Johnson JA, Meisner LF, Jones KL, Dalton S, Stice SL (2005) Preserving the Genetic Integrity of Human Embryonic Stem Cells. *Nat Biotechnol* **23** (1): 19 - 20

Mukhopadhyay D, Dasso M (2007) Modification in Reverse: the SUMO Proteases. *Trends in biochemical sciences* **32** (6): 286 - 295

Muller S, Hoege C, Pyrowolakis G, Jentsch S (2001) SUMO, Ubiquitin's Mysterious Cousin. *Nature reviews.Molecular cell biology* **2** (3): 202 - 210

Nacerddine K, Lehembre F, Bhaumik M, Artus J, Cohen-Tannoudji M, Babinet C, Pandolfi PP, Dejean A (2005) The SUMO Pathway Is Essential for Nuclear Integrity and Chromosome Segregation in Mice. *Developmental cell* **9** (6): 769 - 779

Nagaoka M, Duncan SA (2010) Transcriptional Control of Hepatocyte Differentiation. *Prog Mol Biol Transl.Sci* **97** 79 - 101

Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI (2001) Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. *Science* **292** (5514): 110 - 113

Navas MA, Munoz-Elias EJ, Kim J, Shih D, Stoffel M (1999) Functional Characterization of the MODY1 Gene Mutations HNF4(R127W), HNF4(V255M), and HNF4(E276Q). *Diabetes* **48** (7): 1459 - 1465

Navas MA, Vaisse C, Boger S, Heimesaat M, Kollee LA, Stoffel M (2000) The Human HNF-3 Genes: Cloning, Partial Sequence and Mutation Screening in Patients With Impaired Glucose Homeostasis. *Hum.Hered*. **50** (6): 370 - 381

Ng S, Wu YN, Zhou Y, Toh YE, Ho ZZ, Chia SM, Zhu JH, Mao HQ, Yu H (2005) Optimization of 3-D Hepatocyte Culture by Controlling the Physical and Chemical Properties of the Extra-Cellular Matrices. *Biomaterials* **26** (16): 3153 - 3163

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, ¶ler H, Smith A (1998) Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell* **95** (3): 379 - 391

Niebruegge S, Bauwens CL, Peerani R, Thavandiran N, Masse S, Sevaptisidis E, Nanthakumar K, Woodhouse K, Husain M, Kumacheva E, Zandstra PW (2009) Generation of Human Embryonic Stem Cell-Derived Mesoderm and Cardiac Cells Using Size-Specified Aggregates in an Oxygen-Controlled Bioreactor. *Biotechnol Bioeng* **102** (2): 493 - 507

Nowak M, Hammerschmidt M (2006) Ubc9 Regulates Mitosis and Cell Survival During Zebrafish Development. *Molecular biology of the cell* **17** (12): 5324 - 5336

Nur-E-Kamal, Ahmed I, Kamal J, Schindler M, Meiners S (2006) Three-Dimensional Nanofibrillar Surfaces Promote Self-Renewal in Mouse Embryonic Stem Cells. *Stem Cells* **24** (2): 426 - 433

Ober EA, Verkade H, Field HA, Stainier DYR (2006) Mesodermal Wnt2b Signalling Positively Regulates Liver Specification. *Nature* **442** (7103): 688 - 691

Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, Fraenkel E, Bell GI, Young RA (2004) Control of Pancreas and Liver Gene Expression by HNF Transcription Factors. *Science* **303** (5662): 1378 - 1381

Opazo P, Watabe AM, Grant SGN, O'Dell TJ (2003) Phosphatidylinositol 3-Kinase Regulates the Induction of Long-Term Potentiation Through Extracellular Signal-Related Kinase-Independent Mechanisms. *The Journal of Neuroscience* **23** (9): 3679 - 3688

Parviz F, Matullo C, Garrison WD, Savatski L, Adamson JW, Ning G, Kaestner KH, Rossi JM, Zaret KS, Duncan SA (2003) Hepatocyte Nuclear Factor 4alpha Controls the Development of a Hepatic Epithelium and Liver Morphogenesis. *Nat Genet* **34** (3): 292 - 296

Pedersen TA, Kowenz-Leutz E, Leutz A, Nerlov C (2001) Cooperation Between C/EBPalpha TBP/TFIIB and SWI/SNF Recruiting Domains Is Required for Adipocyte Differentiation. *Genes & development* **15** (23): 3208 - 3216

Peiler G, Bockmann B, Nakhei H, Ryffel GU (2000) Inhibitor of the Tissue-Specific Transcription Factor HNF4, a Potential Regulator in Early Xenopus Development. *Mol Cell Biol* **20** (23): 8676 - 8683

Pera MF, Andrade J, Houssami S, Reubinoff B, Trounson A, Stanley EG, Ward-van Oostwaard D, Mummery C (2004) Regulation of Human Embryonic Stem Cell Differentiation by BMP-2 and Its Antagonist Noggin. *J Cell Sci* **117** (Pt 7): 1269 - 1280

Pera MF, Reubinoff B, Trounson A (2000) Human Embryonic Stem Cells. *Journal of cell science* **113** (1): 5 - 10

Perrimon N, Ni JQ, Perkins L (2010) In Vivo RNAi: Today and Tomorrow. *Cold Spring Harbor Perspectives in Biology* **2** (8):

Peura T, Schaft J, Dumevska B, Stojanov T (2011) Generation of Human Embryonic Stem Cells. *Curr Protoc Stem Cell Biol* **Chapter 1** Unit1A -

Pickett FB, Meeks-Wagner DR (1995) Seeing Double: Appreciating Genetic Redundancy. *Plant Cell* **7** (9): 1347 - 1356

Postovit LM, Seftor EA, Seftor RE, Hendrix MJ (2006) A Three-Dimensional Model to Study the Epigenetic Effects Induced by the Microenvironment of Human Embryonic Stem Cells. *Stem Cells* **24** (3): 501 - 505

Prowse AB, McQuade LR, Bryant KJ, Marcal H, Gray PP (2007) Identification of Potential Pluripotency Determinants for Human Embryonic Stem Cells Following Proteomic Analysis of Human and Mouse Fibroblast Conditioned Media. *J Proteome.Res* **6** (9): 3796 - 3807

Peura TT, Bosman A, Stojanov T (2007) Derivation of Human Embryonic Stem Cell Lines. *Theriogenology* **67** (1): 32 - 42

Prudden J, Pebernard S, Raffa G, Slavin DA, Jefferson J, Perry P, Tainer JA, McGowan CH, Boddy MN (2007) SUMO-targeted ubiquitin ligases in genome stability. *The EMBO Journal* (2007) **26**: 4089 - 4101

Rajala K, Lindroos B, Hussein SM, Lappalainen RS, Pekkanen-Mattila M, Inzunza J, Rozell B, Miettinen S, Narkilahti S, Kerkela E, Aalto-Setala K, Otonkoski T, Suuronen R, Hovatta O, Skottman H (2010) A Defined and Xeno-Free Culture Method Enabling the Establishment of Clinical-Grade Human Embryonic, Induced Pluripotent and Adipose Stem Cells. *PLoS ONE* **5** (4): e10246 -

Rambhatla L, Chiu CP, Kundu P, Peng Y, Carpenter MK (2003) Generation of Hepatocyte-Like Cells From Human Embryonic Stem Cells. *Cell Transplant.* **12** (1): 1 - 11

Rashid ST, Corbineau S, Hannan N, Marciniak SJ, Miranda E, Alexander G, Huang-Doran I, Griffin J, Ahrlund-Richter L, Skepper J, Semple R, Weber A, Lomas DA, Vallier L (2010) Modelling Inherited Metabolic Disorders of the Liver Using Human Induced Pluripotent Stem Cells. *J Clin Invest* **120** (9): 3127 - 3136

Reijo Pera RA, DeJonge C, Bossert N, Yao M, Hwa Yang JY, Asadi NB, Wong W, Wong C, Firpo MT (2009) Gene Expression Profiles of Human Inner Cell Mass Cells and Embryonic Stem Cells. *Differentiation* **78** (1): 18 - 23

Ribeiro A, Archer A, Le Beyec J, Cattin A, Saint-Just S, Pinçon-Raymond M, Chambaz J, Lacasa M, Cardot P (2007) Hepatic Nuclear Factor-4, a Key Transcription Factor at the Crossroads Between Architecture and Function of Epithelia. *Recent Patents Endocr.Metab.Immune Drug Discovery* **1** 166 - 175

Rodin S, Domogatskaya A, Strom S, Hansson EM, Chien KR, Inzunza J, Hovatta O, Tryggvason K (2010) Long-Term Self-Renewal of Human Pluripotent Stem Cells on Human Recombinant Laminin-511. *Nat Biotechnol* **28** (6): 611 - 615

Rodriguez MS, Dargemont C, Hay RT (2001) SUMO-1 Conjugation in Vivo Requires Both a Consensus Modification Motif and Nuclear Targeting. *J Biol Chem* **276** (16): 12654 - 12659

Rodriguez MS, Desterro JM, Lain S, Midgley CA, Lane DP, Hay RT (1999) SUMO-1 Modification Activates the Transcriptional Response of P53. *EMBO J* **18** (22): 6455 - 6461

Rose RB, Endrizzi JA, Cronk JD, Holton J, Alber T (2000) High-Resolution Structure of the HNF-1alpha Dimerization Domain. *Biochemistry* **39** (49): 15062 - 15070

Rossi JM, Dunn NR, Hogan BL, Zaret KS (2001) Distinct Mesodermal Signals, Including BMPs From the Septum Transversum Mesenchyme, Are Required in Combination for Hepatogenesis From the Endoderm. *Genes Dev* **15** (15): 1998 - 2009

Runge D, Runge DM, Drenning SD, Bowen WC, Jr., Grandis JR, Michalopoulos GK (1998) Growth and Differentiation of Rat Hepatocytes: Changes in Transcription Factors HNF-3, HNF-4, STAT-3, and STAT-5. *Biochem Biophys Res Commun* **250** (3): 762 - 768

Ryffel GU (2001) Mutations in the Human Genes Encoding the Transcription Factors of the Hepatocyte Nuclear Factor (HNF)1 and HNF4 Families: Functional and Pathological Consequences. *J Mol Endocrinol* **27** (1): 11 - 29

Salvesen GS (2002) Caspases: Opening the Boxes and Interpreting the Arrows. *Cell Death Differ* **9** (1): 3 - 5

Sato Y, Miyake K, Kaneoka H, Iijima S (2006) Sumoylation of CCAAT/Enhancer-Binding Protein Alpha and Its Functional Roles in Hepatocyte Differentiation. *The Journal of biological chemistry* **281** (31): 21629 - 21639

Schwarts RE (2005) Defined Conditions for Development of Functional Hepatic Cells From Human Embryonic Stem Cells. *Stem cells and development 14* 643 - 655

Selden C, Mellor N, Rees M, Laurson J, Kirwan M, Escors D, Collins M, Hodgson H (2007) Growth Factors Improve Gene Expression After Lentiviral Transduction in Human Adult and Fetal Hepatocytes. *The Journal of Gene Medicine* **9** (2): 67 - 76

Seppen J, Rijnberg M, Cooreman MP, Oude Elferink RPJ (2002) Lentiviral Vectors for Efficient Transduction of Isolated Primary Quiescent Hepatocytes. *Journal of hepatology* **36** (4): 459 - 465

Serls AE, Doherty S, Parvatiyar P, Wells JM, Deutsch GH (2005) Different Thresholds of Fibroblast Growth Factors Pattern the Ventral Foregut into Liver and Lung. *Development (Cambridge, England)* **132** (1): 35 - 47

Sharma R, Greenhough S, Medine CN, Hay DC (2010) Three-Dimensional Culture of Human Embryonic Stem Cell Derived Hepatic Endoderm and Its Role in Bioartificial Liver Construction. *Journal of biomedicine & biotechnology* **2010** 236147 -

Shen LN, Geoffroy M, Jaffray EG, Hay RT (2009) Characterization of SENP7, a SUMO-2/3-Specific Isopeptidase. *Biochemical Journal* **421** (2): 223 - 230

Sheng W, Yan H, Rausa FM, III, Costa RH, Liao X (2004) Structure of the Hepatocyte Nuclear Factor 6alpha and Its Interaction With DNA. *J Biol Chem* **279** (32): 33928 - 33936

Shin D, Shin CH, Tucker J, Ober EA, Rentzsch F, Poss KD, Hammerschmidt M, Mullins MC, Stainier DYR (2007) Bmp and Fgf Signalling Are Essential for Liver Specification in Zebrafish. *Development (Cambridge, England)* **134** (11): 2041 - 2050

Shiojiri N (1997) Development and Differentiation of Bile Ducts in the Mammalian Liver. *Microsc.Res Tech.* **39** (4): 328 - 335

Shirahashi H, Wu J, Yamamoto N, Catana A, Wege H, Wager B, Okita K, Zern MA (2004) Differentiation of Human and Mouse Embryonic Stem Cells Along a Hepatocyte Lineage. *Cell Transplant.* **13** (3): 197 - 211

Sladek FM, Ruse MD, Jr., Nepomuceno L, Huang SM, Stallcup MR (1999) Modulation of Transcriptional Activation and Coactivator Interaction by a Splicing Variation in the F Domain of Nuclear Receptor Hepatocyte Nuclear Factor 4alpha1. *Mol Cell Biol* **19** (10): 6509 - 6522

Sladek FM, Zhong WM, Lai E, Darnell JE (1990) Liver-Enriched Transcription Factor HNF-4 Is a Novel Member of the Steroid Hormone Receptor Superfamily. *Genes & development* **4** (12b): 2353 - 2365

Sladek R, Giguere V (2000) Orphan Nuclear Receptors: an Emerging Family of Metabolic Regulators. *Adv.Pharmacol* **47** 23 - 87

Snykers S, De Kock J, Rogiers V, Vanhaecke T (2009) In Vitro Differentiation of Embryonic and Adult Stem Cells into Hepatocytes: State of the Art. *Stem Cells* **27** (3): 577 - 605

Sokal EM (2011) From Hepatocytes to Stem and Progenitor Cells for Liver Regenerative Medicine: Advances and Clinical Perspectives. *Cell Prolif* **44 Suppl 1** 39 - 43

Sonenberg N, Hinnebusch AG (2009) Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136** (4): 731 - 745

Soutoglou E, Katrakili N, Talianidis I (2000) Acetylation Regulates Transcription Factor Activity at Multiple Levels. *Mol Cell* **5** (4): 745 - 751

Spath GF, Weiss MC (1997) Hepatocyte Nuclear Factor 4 Expression Overcomes Repression of the Hepatic Phenotype in Dedifferentiated Hepatoma Cells. *Mol Cell Biol* **17** (4): 1913 - 1922

Spiro RG (2002) Protein Glycosylation: Nature, Distribution, Enzymatic Formation, and Disease Implications of Glycopeptide Bonds. *Glycobiology* **12** (4): 43R - 56R

Stanger BZ (2008) HNF4A and Diabetes. *Diabetes* **57** (6): 1461 - 1462

Stern CD (2005) Neural Induction: Old Problem, New Findings, Yet More Questions. *Development* **132** (9): 2007 - 2021

Stoffel M, Duncan SA (1997) The Maturity-Onset Diabetes of the Young (MODY1) Transcription Factor HNF4alpha Regulates Expression of Genes Required for Glucose Transport and Metabolism. *Proc Natl Acad Sci U S A* **94** (24): 13209 - 13214

Strulovici Y, Leopold PL, O'Connor TP, Pergolizzi RG, Crystal RG (2007) Human Embryonic Stem Cells and Gene Therapy. *Mol Ther* **15** (5): 850 - 866

Sugii S, Kida Y, Kawamura T, Suzuki J, Vassena R, Yin YQ, Lutz MK, Berggren WT, Izpisua Belmonte JC, Evans RM (2010) Human and Mouse Adipose-Derived Cells Support Feeder-Independent Induction of Pluripotent Stem Cells. *Proc Natl Acad Sci U S A* **107** (8): 3558 - 3563

Suksaweang S, Lin CM, Jiang TX, Hughes MW, Widelitz RB, Chuong CM (2004) Morphogenesis of Chicken Liver: Identification of Localized Growth Zones and the Role of Beta-Catenin/Wnt in Size Regulation. *Dev Biol* **266** (1): 109 - 122

Sun H, Leverson JD, Hunter T (2007) Conserved Function of RNF4 Family Proteins in Eukaryotes: Targeting a Ubiquitin Ligase to SUMOylated Proteins. *EMBO J* **26** (18): 4102 - 4112

Sun K, Montana V, Chellappa K, Brelivet Y, Moras D, Maeda Y, Parpura V, Paschal BM, Sladek FM (2007) Phosphorylation of a Conserved Serine in the Deoxyribonucleic Acid Binding Domain of Nuclear Receptors Alters Intracellular Localization. *Mol Endocrinol* **21** (6): 1297 - 1311

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of Pluripotent Stem Cells From Adult Human Fibroblasts by Defined Factors. *Cell* **131** (5): 861 - 872

Takebe Y, Seiki M, Fujisawa J, Hoy P, Yokota K, Arai K, Yoshida M, Arai N (1988) SR Alpha Promoter: an Efficient and Versatile Mammalian CDNA Expression System Composed of the Simian Virus 40 Early Promoter and the R-U5 Segment of Human T-Cell Leukemia Virus Type 1 Long Terminal Repeat. *Molecular and cellular biology* **8** (1): 466 - 472

Tam PPL, Kanai-Azuma M, Kanai Y (2003) Early Endoderm Development in Vertebrates: Lineage Differentiation and Morphogenetic Function. *Current opinion in genetics & development* **13** (4): 393 - 400

Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C, Umezu H, Hirayama Y, Midorikawa Y, Hippo Y, Watanabe A, Uchiyama Y, Hasegawa G, Reid P, Aburatani H, Hamakubo T, Sakai J, Naito M, Kodama T (2006) Dysregulated Expression of P1 and P2 Promoter-Driven Hepatocyte Nuclear Factor-4alpha in the Pathogenesis of Human Cancer. *J Pathol* **208** (5): 662 - 672

Tatham MH, Kim S, Yu B, Jaffray E, Song J, Zheng J, Rodriguez MS, Hay RT, Chen Y (2003) Role of an N-Terminal Site of Ubc9 in SUMO-1, -2, and -3 Binding and Conjugation. *Biochemistry* **42** (33): 9959 - 9969

Tatham MH, Geoffroy MC, Shen L, Plechanovova A, Hattersley N, Jaffray EG, Palvimo JJ, Hay RT (2008) RNF4 Is a Poly-SUMO-Specific E3 Ubiquitin Ligase Required for Arsenic-Induced PML Degradation. *Nature cell biology* **10** (5): 538 - 546

Tavakoli T, Xu X, Derby E, Serebryakova Y, Reid Y, Rao MS, Mattson MP, Ma W (2009) Self-Renewal and Differentiation Capabilities Are Variable Between Human Embryonic Stem Cell Lines I3, I6 and BG01V. *BMC Cell Biol* **10** 44 -

Teija TP, Alexis B, Tomas S (2007) Derivation of Human Embryonic Stem Cell Lines. *Theriogenology* (67): 32 - 42

Terai S, Sakaida I, Yamamoto N, Omori K, Watanabe T, Ohata S, Katada T, Miyamoto K, Shinoda K, Nishina H, Okita K (2003) An In Vivo Model for Monitoring Trans-Differentiation of Bone Marrow Cells into Functional Hepatocytes. *Journal of biochemistry* **134** (4): 551 - 558

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD (2007) New Cell Lines From Mouse Epiblast Share Defining Features With Human Embryonic Stem Cells. *Nature* **448** (7150): 196 - 199

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM (1998) Embryonic Stem Cell Lines Derived From Human Blastocysts. *Science (New York, N.Y.)* **282** (5391): 1145 - 1147

Tilles AW, Berthiaume F, Yarmush ML, Toner M (2002) Critical Issues in Bioartificial Liver Development. *Technol.Health Care* **10** (3-4): 177 - 186

Tomas S (2007) A Large Scale Proteomic Analysis of Human Embryonic Stem Cells. *BMC Genomics* **8** (478):

Torres-Padilla ME, Sladek FM, Weiss MC (2002) Developmentally Regulated N-Terminal Variants of the Nuclear Receptor Hepatocyte Nuclear Factor 4Iñ Mediate Multiple Interactions Through Coactivator and Corepressor-Histone Deacetylase Complexes. *Journal of Biological Chemistry* **277** (47): 44677 - 44687

Trichas G, Begbie J, Srinivas S (2008) Use of the Viral 2A Peptide for Bicistronic Expression in Transgenic Mice. *BMC Biology* **6** (1): 40 -

Trounson A (2006) The Production and Directed Differentiation of Human Embryonic Stem Cells. *Endocr Rev* **27** (2): 208 - 219

Twig G, Hyde B, Shirihai OS (2008) Mitochondrial Fusion, Fission and Autophagy As a Quality Control Axis: the Bioenergetic View. *Biochimica et biophysica acta* **1777** (9): 1092 - 1097

Ulrich HD (2008) The Fast-Growing Business of SUMO Chains. *Mol Cell* **32** (3): 301 - 305

Ulrich HD (2009) The SUMO System: an Overview. *Methods in molecular biology (Clifton, N.J.)* **497** 3 - 16

Uzunova K, Gottsche K, Miteva M, Weisshaar SR, Glanemann C, Schnellhardt M, Niessen M, Scheel H, Hofmann K, Johnson ES, Praefcke GJ, Dohmen RJ (2007) Ubiquitin-Dependent Proteolytic Control of SUMO Conjugates. *J Biol Chem* **282** (47): 34167 - 34175

Vaillancourt C , Lafond J (2009) Human Embryogenesis: Overview. *Methods Mol Biol* **550** 3 - 7

van Hagen M, Overmeer RM, Abolvardi SS, Vertegaal ACO (2010) RNF4 and VHL Regulate the Proteasomal Degradation of SUMO-Conjugated Hypoxia-Inducible Factor-2Î±. *Nucleic acids research* **38** (6): 1922 - 1931

Van Hoof D, Yoz J, Braam SR, Pinkse MWH, Linding R, Heck AJR, Mummery CL, Krijgsveld J (2009) Phosphorylation Dynamics During Early Differentiation of Human Embryonic Stem Cells. *Cell stem cell* **5** (2): 214 - 226

Vazin T , Freed WJ (2010) Human Embryonic Stem Cells: Derivation, Culture, and Differentiation: a Review. *Restor.Neurol.Neurosci.* **28** (4): 589 - 603

Vessey CJ , de la Hall PM (2001) Hepatic Stem Cells: a Review. *Pathology* **33** (2): 130 - 141

Villa-Diaz LG, Nandivada H, Ding J, Nogueira-de-Souza NC, Krebsbach PH, O'Shea KS, Lahann J, Smith GD (2010) Synthetic Polymer Coatings for Long-Term Growth of Human Embryonic Stem Cells. *Nat Biotechnol* **28** (6): 581 - 583

Viollet B, Kahn A, Raymondjean M (1997) Protein Kinase A-Dependent Phosphorylation Modulates DNA-Binding Activity of Hepatocyte Nuclear Factor 4. *Mol Cell Biol* **17** (8): 4208 - 4219

Waby JS, Bingle CD, Corfe BM (2008) Post-Translational Control of Sp-Family Transcription Factors. *Current genomics* **9** (5): 301 - 311

Waldman T, Kinzler KW, Vogelstein B (1995) P21 Is Necessary for the P53-Mediated G1 Arrest in Human Cancer Cells. *Cancer Res* **55** (22): 5187 - 5190

Walsh CT, Garneau-Tsodikova S, Gatto GJ (2005) Protein Posttranslational Modifications: the Chemistry of Proteome Diversifications. *Angewandte Chemie (International ed.in English)* **44** (45): 7342 - 7372

Wang G, Zhang H, Zhao Y, Li J, Cai J, Wang P, Meng S, Feng J, Miao C, Ding M, Li D, Deng H (2005) Noggin and BFGF Cooperate to Maintain the Pluripotency of Human Embryonic Stem Cells in the Absence of Feeder Layers. *Biochem Biophys Res Commun* **330** (3): 934 - 942

Wang Y, Susando T, Lei X, Anene-Nzelu C, Zhou H, Liang LH, Yu H (2010) Current Development of Bioreactors for Extracorporeal Bioartificial Liver (Review). *Biointerphases.* **5** (3): FA116 - FA131

Watt AJ, Garrison WD, Duncan SA (2003) HNF4: a Central Regulator of Hepatocyte Differentiation and Function. *Hepatology* **37** (6): 1249 - 1253

Watt AJ, Zhao R, Li J, Duncan SA (2007) Development of the Mammalian Liver and Ventral Pancreas Is Dependent on GATA4. *BMC Dev Biol* **7** 37 -

Watt AJ, Garrison WD, Duncan SA (2003) HNF4: A Central Regulator of Hepatocyte Differentiation and Function. *Hepatology* **37** (6): 1249 - 1253

Wege H, Chui MS, Le HT, Strom SC, Zern MA (2003) In Vitro Expansion of Human Hepatocytes Is Restricted by Telomere-Dependent Replicative Aging. *Cell transplantation* **12** (8): 897 - 906

Wei F, ler HR, Atchison ML (2007) Sumoylation of Oct4 Enhances Its Stability, DNA Binding, and Transactivation. *The Journal of biological chemistry* **282** (29): 21551 - 21560

Wisely GB, Miller AB, Davis RG, Thornquest AD, Jr., Johnson R, Spitzer T, Sefler A, Shearer B, Moore JT, Miller AB, Willson TM, Williams SP (2002) Hepatocyte Nuclear Factor 4 Is a Transcription Factor That Constitutively Binds Fatty Acids. *Structure* **10** (9): 1225 - 1234

Wong PK, Campbell IK, Egan PJ, Ernst M, Wicks IP (2003) The Role of the Interleukin-6 Family of Cytokines in Inflammatory Arthritis and Bone Turnover. *Arthritis Rheum.* **48** (5): 1177 - 1189

Wortham M, Czerwinski M, He L, Parkinson A, Wan YJ (2007) Expression of Constitutive Androstane Receptor, Hepatic Nuclear Factor 4 Alpha, and P450 Oxidoreductase Genes Determines Interindividual Variability in Basal Expression and Activity of a Broad Scope of Xenobiotic Metabolism Genes in the Human Liver. *Drug Metab Dispos.* **35** (9): 1700 - 1710

Xiao Q, Luo Z, Pepe AE, Margariti A, Zeng L, Xu Q (2008) Embryonic Stem Cell Differentiation into Smooth Muscle Cells Is Mediated by Nox4-Produced H2O2. *Am J Physiol Cell Physiol*

Xu C, Inokuma MS, Denham J (2001) Feeder Free Growth of Undifferentiated Human Embryonic Stem Cells. *Nature Biotechnology* **19** 971 - 974

Xu C, Inokuma MS, Denham J, Golds K, Kundu P, Gold JD, Carpenter MK (2001) Feeder-Free Growth of Undifferentiated Human Embryonic Stem Cells. *Nat Biotech* **19** (10): 971 - 974

Xu Y, Zhu X, Hahm HS, Wei W, Hao E, Hayek A, Ding S (2010) Revealing a Core Signalling Regulatory Mechanism for Pluripotent Stem Cell Survival and Self-Renewal by Small Molecules. *Proceedings of the National Academy of Sciences* **107** (18): 8129 - 8134

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI (1996) Mutations in the Hepatocyte Nuclear Factor-4[Alpha] Gene in Maturity-Onset Diabetes of the Young (MODY1). *Nature* **384** (6608): 458 - 460

Yao F, Svensjo T, Winkler T, Lu M, Eriksson C, Eriksson E (1998) Tetracycline Repressor, TetR, Rather Than the TetR-Mammalian Cell Transcription Factor Fusion Derivatives, Regulates Inducible
Gene Expression in Mammalian Cells. *Human Gene Therapy* **9** 1939 - 1850

Yin C, Ying L, Zhang PC, Zhuo RX, Kang ET, Leong KW, Mao HQ (2003) High Density of Immobilized Galactose Ligand Enhances Hepatocyte Attachment and Function. *J Biomed Mater.Res A* **67** (4): 1093 - 1104

Ying QL, Nichols J, Chambers I, Smith A (2003) BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration With STAT3. *Cell* **115** (3): 281 - 292

Yuan X, Ta TC, Lin M, Evans JR, Dong Y, Bolotin E, Sherman MA, Forman BM, Sladek FM (2009) Identification of an Endogenous Ligand Bound to a Native Orphan Nuclear Receptor. *PLoS ONE* **4** (5): e5609 -

Zaret KS (2000) Liver Specification and Early Morphogenesis. *Mech Dev* **92** (1): 83 - 88

Zaret KS (2001) Hepatocyte Differentiation: From the Endoderm and Beyond. *Current opinion in genetics & development* **11** (5): 568 - 574

Zhang F, Frost AR, Blundell MP, Bales O, Antoniou MN, Thrasher AJ (2010) A Ubiquitous Chromatin Opening Element (UCOE) Confers Resistance to DNA Methylation-Mediated Silencing of Lentiviral Vectors. *Mol Ther* **18** (9): 1640 - 1649

Zhao R, Duncan SA (2005) Embryonic Development of the Liver. *Hepatology* **41** (5): 956 - 967

Zunino R, Schauss A, Rippstein P, Andrade-Navarro M, McBride HM (2007) The SUMO Protease SENP5 Is Required to Maintain Mitochondrial Morphology and Function. *Journal of cell science* **120** (Pt 7): 1178 - 1188

SUPPLEMENTARY INFORMATION

Supplementary Figure 1 – HNF4 α Splice Variants Primer Binding Sites
The diagram denotes the annealing positions of the specific HNF4 α splice variant primers in RT-PCR. A – HNF4 α Variant 1, B – HNF4 α Variant 2, C – HNF4 α Variant 3, D – HNF4 α Variant 4, E – HNF4 α Variant 5 and F – HNF4 α Variant 6.